THE EVOLUTION OF INVASIVENESS IN POECILIID FISHES: INSIGHTS FROM LIFE HISTORY AND GENOMICS

By

NICHOLAS JOSEPH TROENDLE

(Under the direction of Rodney Mauricio)

ABSTRACT

Understanding how a species becomes invasive and knowing the traits that promote invasiveness are vital for developing effective management strategies. With respect to invasion biology, life history traits are understudied in vertebrate taxa. In this dissertation, I used museum collections to investigate the connection between life history traits and invasiveness in live-bearing fishes from the family Poeciliidae. Since one Poeciliid fish, *Gambusia affinis*, (the Western Mosquitofish) is an important model species for invasion, I completed a draft genome of this species as a community resource. I used this genetic information to reconstruct the historical invasion route of *Gambusia affinis* from the southeastern United States into east Asia.

In Chapter 2, I compared 11 invasive species from the family Poeciliidae to 11 closely related, non-invasive species. An investigation of life history traits showed that invasive species have significantly more offspring but that those offspring are smaller than in non-invasive species. This tradeoff could lead to rapid rates of increase in invasive species allowing them to quickly get a foothold and expand when introduced to a new location.

In Chapter 3, I examined life history traits of eleven invasive Poeciliid species comparing their native and invasive ranges in order to determine if life history traits shifted as a result of becoming invasive. My investigation showed that in their invasive range invasive species displayed life history strategies that increased their population growth rate. Body size, fecundity, and reproductive allotment all increased in the invasive range without sacrificing offspring size. This suggests that when a species colonizes a new environment there are shifts in life history that help them to become established more quickly.

In Chapter 4, I assembled a reference genome for *Gambusia affinis* using Illumina short read sequencing of traditional paired end libraries in conjunction with the new Chicago Libraries. The reference genome has high contiguity and coverage, with N50 contig and N50 scaffold lengths of 17.6Kb and 6.65Mb, respectively, and total estimated coverage of 55X. I then annotated the genome and compared its quality to three other fish genomes to ensure that the assembly was of high quality.

In Chapter 5, I used RadSeq data to produce a suite of 4405 SNPs, mapped to the *Gambusia affinis* genome, for 12 populations of the invasive Western Mosquitofish *G. affinis* along its presumed invasion route from the southeastern United States to east Asia. I then used these markers to perform statistical and phylogenetic analysis in order to assess the genetic diversity in the invasive range and to determine the likelihood of multiple introductions. I found higher levels of genetic diversity in parts of the invasive range than would be expected from serial introductions over such a short time period. This is likely due to multiple introductions, possibly of large numbers of individuals.

THE EVOLUTION OF INVASIVENESS IN POECILIID FISHES: INSIGHTS FROM LIFE HISTORY AND GENOMICS

By

NICHOLAS JOSEPH TROENDLE

B.S., Pepperdine University, 2009

M.S., Texas A&M University, 2012

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

THE EVOLUTION OF INVASIVENESS IN POECILIID FISHES: INSIGHTS FROM LIFE HISTORY AND GENOMICS

By

NICHOLAS JOSEPH TROENDLE

Major Professor: Rodney Mauricio

Committee: Allen Moore
Travis Glen
Byron Freeman
Brian Langerhans

Electronic Version Approved:

Suzanne Barbour
Dean of Graduate School
The University of Georgia
August 2016

ACKNOWLEDGEMENTS

The National Science Foundation Partnerships in International Research and Education Program (Grant No. OISE 0730218) provided the primary funding for this research. This research was also supported by the Georgia Genomics Facility, at the University of Georgia. Also, with the support of Dovetail Genomics, who allowed us to be part of their Alpha testing group to produce our genome assembly.

I am thankful to my advisor, Rodney Mauricio, who allowed me to join his lab and develop my own research project despite the fact that I did not want to work on plants. Each of my committee members (Allen Moore, Travis Glenn, Byron Freeman, and Brian Langerhans) has provided invaluable advice and assistance throughout the course of this project.

I am grateful to the Mauricio lab, both past and present members for their assistance. Jared Lee provided collection locations, equipment, samples and advice for my research. Sandra Hoffberg assisted me in wading through the Stacks pipeline and was always being willing to talk things through. I would also like to thank Ousman Mahmud and Domitille Chalopin for their invaluable assistance in annotating the genome.

My research could not have been completed without the use of extensive specimen collections and I am thankful for the people and institutions that provided these: Douglas Nelson (University of Michigan Museum of Zoology), Adam Cohen (Texas Natural History Collections), Mark Sabaj Perez (Department of Ichthyology Academy of Natural Sciences), Tulane University Museum of Natural History, Lynne

Parenti (Smithsonian Institution National Museum of Natural History), David Boyd (Division of Ichthyology, Florida Museum of Natural History).

I would like to thank my children for always providing the time away from work that I needed to keep my sanity. Finally, I would like to thank my loving and supportive wife Stephanie, who has always believed that I could get through this and has been my partner in every part of the journey, from helping me make fish breeding baskets to driving 2500 miles with buckets of fish in the back seat. Without you none of this would have been possible and I can't wait to see what God has in store for us next.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Invasive species have been an issue of growing concern for the past several decades. This has not, however, prevented species from being transplanted both intentionally and accidentally. Today, it is estimated that there are approximately 50,000 non-native species in the United States alone (Pimentel, 2005). However, not every species that is non-native is considered invasive. In order to be considered invasive a species usually needs to meet two criteria. First, it must be non-native and second it must do harm to other species, the ecosystem or human health (Executive Order 13112). Of the non-native species present in the United States, 4,300 are considered invasive (Corn *et al.* 1999).

Invasive species have significant biological impacts on native species and environments (Elton, 1958; Charles & Dukes, 2007). Invasive species can cause environmental change worldwide leading to a loss of biodiversity, degrading of environmental health and economic costs (Vitousek *et al.* 1996; Didham *et al.* 2005). There are many ways that invasive species affect their new environments. They often cause the decline of native species through predation or direct competition (Rehage *et al.* 2005). In other cases, hybridization between invasive and native species can lead to decreased fitness and dilution of the native phenotypes eventually resulting in a loss of any true natives (Huxel, 1998; Lee, 2002). In some cases, these factors, or a combination of them, and the introduction of novel diseases or parasites, may lead to extinction of native flora and fauna, resulting in an overall decay of earth's biodiversity (Cox 2004; Dextrase & Mandrake, 2006).

Despite their devastating negative environmental impacts, invasive species have provided unique opportunities for ecological and evolutionary biological research. As they are subjected to novel environments and selection pressures, invasive species are natural experiments in evolution (Sakai *et al.* 2001; Huey *et al.* 2005). This means that we can observe the process of evolution as it happens in these invasive species.

Over 40 years ago Baker (1974), proposed the concept of an ideal weed in which he outlined the traits that a perfect invasive species should have. According to Baker, the ideal weed would grow rapidly, have high fecundity, and short generation time. It would also have diverse dispersal methods, be self-compatible and have a wide range of ecological tolerances. As Baker observed, this is a list for the ideal weed; no real species has all of the characteristics that he outlined, although most invasive species have a subset of them. Despite the amount of research invested, we still do not fully understand what traits make a species a good invader.

Most of the traits on Baker's list have to do with the growth and reproduction of the species, which we call life history traits. Life history traits are the vital rates of an organism. Traits like age at sexual maturity, number and size of offspring, reproductive duration and senescence all directly affect an organism's fitness and the intrinsic rates of increase of a population (Stearns, 1976; Reznick *et al.* 1996). As such, they are often under strong selection. Life history traits are generally constrained by trade-offs either due to limited energy or time (Roff, 2002). For example, given a finite number of resources an organism must "decide" how to divide up those resources between growth and reproduction. This leads to different suites of life history traits in different environments as different strategies optimize fitness under different ecological conditions. Because of this, it has been speculated that life history traits might be one of the driving forces behind successful invasion (Sakai *et al.* 2001; Castro-Diez *et al.* 2011;

Kolar & Lodge, 2002). While the link between life history and invasion has been studied, vertebrate taxa are highly underrepresented with the majority of studies focusing on plant species. There is no consensus as to what suite of life history traits make an invader successful. However, studies that investigate many species tend to make stronger conclusions than those that simply look at a few (Olden *et al.* 2006; Rejmanek & Richardson, 1996; Rosecchi *et al.* 2001; Radford & Cousens, 2000). For example, Rejmanek and Richardson (1996) were able to successfully predict invasive ability in pine trees using life history traits from 24 species as predictors. They were then able to use their predictors to correctly predict the invasive ability of non-pine species as well.

One of the reasons that the vast majority of invasive research is done on plant species is because they are by far the most plentiful group of invasive species. However, if you want to study vertebrates the best group to study would be fish. Of the 100 worst invasive species list produced by the Invasive Species Specialist Group, 8 are fish species (Lowe *et al.* 2000). While this may not seem like a large proportion of species fish make up nearly 1/3 of the vertebrate species on that list. Fish are by far the most invasive group of vertebrates. Pimentel (2011), for example found that out of the 138 invasive vertebrates in Brazil 79.7% of them were fish species. The abundance and diversity of invasive fishes makes them ideal for such investigations.

The family of live bearing fishes, Poeciliidae, provides an excellent system for studying invasion-associated life history. Due to intentional introductions for mosquito control and their popularity in the aquarium trade, there are multiple invasive species in the family (Sakai *et al.* 2001). These invasive species can be compared to closely related non-invasive species. Also, because they are live-bearing fishes and females become visibly gravid it is relatively simple to

pick out pregnant female fish to be dissected. Also Poeciliidae have been used as a model of life history evolution as it is possible to count and stage each embryo (Reznick & Endler, 1982).

It is also important to understand the shifts in life history traits that are likely to occur when a species becomes invasive. Since invasive species can be considered experiments in evolution, we cannot expect their characteristics to remain the same. It is also important to realize that simply looking at invasive and non-invasive species may not show us the full picture. For example, an invasive species may have similar traits to a non-invasive species when the invasive species is in its native range. However, it may change when it is introduced to a new environment. If this were the case, we would not see the traits that promote invasion by observing only the native range. By observing changes in the invasive range of a species, we can investigate evolution in a natural system in real time. Invasive species often seem to change relatively quickly; sometimes we see rapid shifts due to phenotypic plasticity, environmentally caused variation within a genotype. In other cases, adaptation proceeds rapidly.

One important part of understanding invasive species is knowing the amount of genetic diversity in its invasive range, as this is likely to affect the evolutionary potential of the species in that range (Roman and Darling, 2007). When a species is introduced the introduced range is expected to have reduced genetic variation (Dlugosch and Parker, 2008). This should result in reduced evolutionary potential however many invasive species still seem to thrive around the world (Tsutsui *et al.* 2000; Frankham, 2005), which is known as the 'genetic paradox.' It has recently been discovered that while we expect to see a decrease in genetic diversity in invasive species, this is not as universal as has been expected (Roman and Darling, 2007). In some species, genetic diversity has actually increased in the invasive range (Kolbe *et al.* 2004; Genton

et al. 2005; Lindholm *et al.* 2005). This unexpected increase in diversity has been attributed to multiple introductions, generally of a large number of individuals (Dlugosh and Parker, 2008).

Although native to the southeastern United States, the western mosquitofish, *Gambusia affinis,* was broadly introduced throughout the world in the early 20th century as a means of mosquito control in areas where malaria and yellow fever were common (Krumholz 1948; Pyke 2008). While their efficacy is a subject of much debate, *G. affinis,* and its sister species *G. holbrooki*, have become extremely widespread and today are present in over 50 countries (Pyke, 2008; Lowe *et al.* 2000). In this case, these two species have become the most widespread freshwater fish in the world in just under 100 years.

In the following chapters, I investigate the family of livebearing fish Poeciliidae to determine what life history traits might be associated with invasion. I also use genomic data to assess the genetic diversity of *Gambusia affinis* in the invasive range. In Chapter 2, I compare the life history traits of eleven invasive Poeciliid species to eleven, closely related, non-invasive species. This study examines the patterns of life history evolution in invasive species. While similar studies have been performed they have primarily been focused on plant systems and an extensive vertebrate study is needed.

In Chapter 3, I use the eleven invasive species from Chapter 2 to investigate how the act of invasion affects life history. This study compares the life history traits of invasive, Poecillid fishes in their native and invasive range. This provides the second half of the picture from Chapter 2. Since invasive species often evolve rapidly and life history traits can be under extreme selection examining only the native range of an invasive species provides an incomplete picture of how life history affects invasive ability.

In Chapter 4, I outline the creation of a reference genome for the invasive, Poeciliid fish *Gambusia affinis*. As *G. affinis* is one of the most invasive fish species in the world it has been highly studied but the types of research have been limited by the lack of a high quality reference genome. The genome that we have constructed could allow researchers to obtain a greater understanding of the extreme invasiveness of this species.

In Chapter 5, I use the genome of the Mosquitofish (*G. affinis*), that we generated in Chapter 4, along with RadSeq data to generate a set of 4405 SNPs. I use these loci to conduct phylogenetic and statistical analyses to examine the amount of genetic diversity contained in the invasive range to evaluate the likelihood of multiple introductions to some parts of the invasive range.

These four chapters further my goal of better understanding how Poeciliid fishes become invasive and what types of traits promote the invasiveness of some species while other species remain isolated to their native ranges. Using Poeciliids I demonstrate the potential effects that life history can have on invasive species. I also show the importance of sampling invasive species from both their native and invasive range when examining traits that may be under high levels of selection.

References

Baker, H.G. 1974. The evolution of weeds. Annual Review of Ecology and Systematics **5**: 1-24.

Castro-Díez, P., O. Godoy, A. Saldana, and D.M. Richardson. 2011. Predicting invasiveness of Australian acacias on the basis of their native climatic affinities, life history traits and human use. Diversity and Distributions **17**: 934-945.

Charles, H., and J.S. Dukes. 2007. Impacts of invasive species on ecosystem services. Pp. 217-237 *in*: W. Nentwick (editor). Biological Invasions. Springer, Berlin, Germany.

Corn, L.C., E.H. Buck, J. Rawson, and E. Fischer. 1999. Harmful non-native species: Issues for Congress. Congressional Research Service Issue Brief, RL30123.

Cox G.W. 2004. Alien Species and Evolution. The Evolutionary Ecology of Exotic Plants, Animals, Microbes and Interacting Native Species. Island Press, Washington DC.

Dextrase, A.J., and N.E. Mandrak. 2006. Impacts of alien invasive species on freshwater fauna at risk in Canada. Biological Invasions **8**: 13-24.

Didham, R.K., J.M. Tylianakis, M.A. Hutchinson, R.M. Ewers, and N.J. Gemmell. 2005. Are invasive species the drivers of ecological change? Trends in Ecology and Evolution **20**: 470-474.

Elton C.S. 1958. The ecology of invasions by animals and plants. Progress in Physical Geography **31**: 659-666.

Huey, R.B., G.W. Gilchrist, and A.P. Hendry. 2005. Using invasive species to study evolution. Pp. 138-164 *in*: D.F. Sax, J.J. Stachowicz, and S.D. Gaines (editors). Species Invasions: Insights to Ecology, Evolution and Biogeography. Sinauer Associates, Sunderland, MA.

Huxel G.R. 1999. Rapid displacement of native species by invasive species: effects of hybridization. Biological Conservation **89**: 143-152.

Kolar, C.S. and D.M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. Science **298**: 1233-1236.

Krumholz, L. 1948. Reproduction in the western mosquitofish, *Gambusia affinis affinis* (Baird & Girard), and its use in mosquito control. Ecological Monographs **18**: 1-43.

Lee, C.E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution **17**: 386-391.

Lindholm, A.K., F. Breden, H.J. Alexander, W.K. Chan, S.G. Thakurta, and R. Brooks. 2005. Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. Molecular Ecology **14**: 3671-3682.

Lowe S., B.M. Boudjelas, S.M. De Poorter. 2000. 100 of the World's Worst Invasive Alien Species. The Invasive Species Specialist Group. **11**: 1-13.

Olden J.D., P.N.L. Bestgen R. Kevin. 2006. Life-history strategies predict fish invasions and extirpations in the Colorado river basin. Ecological Monographs **76**: 25-40.

Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annual Review of Ecological Systems **32**: 305-332.

Pimentel, D., R. Zuniga, and D. Morrison. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics **52**: 273-288.

Pimentel, D. 2011. Biological invasions: Economic and environmental costs of alien plant, animal, and microbe species. CRC Press, Boca Raton, Fl.

Pyke, G.H. 2008. Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. Annual Review of Ecology, Evolution, and Systematics **39**: 171-191.

Radford, I.J. and R.D. Cousens. 2000. Invasiveness and comparative life-history traits of exotic and indigenous *Senecio* species in Australia. Oecologia **125**: 531-542.

Rehage, J.S., B.K. Barnett, and A. Sih. 2005. Foraging behavior and invasiveness: do invasive *Gambusia* exhibit higher feeding rates and broader diets than their noninvasive relatives? Ecology of Freshwater Fish **14**: 352-360.

Rejmanek, M., and D.M. Richardson. 1996. What attributes make some plant species more invasive? Ecology **77**: 1655-1661.

Reznick, D. and J.A. Endler. 1982. The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). Evolution **36**: 160-177.

Reznick, D.N., F.H. Rodd, and M. Cardenas. 1996. Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. The American Naturalist **147**: 319-338.

Roff, D.A. 2002. Life History Evolution. Sinauer Associates, Sunderland, MA.

Rosecchi, E., F. Thomas, and A.J. Crivelli. 2001. Can life-history traits predict the fate of introduced species? A case study on two cyprinid fish in southern France. Freshwater Biology **46**: 845-853.

Sakai A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, With A. Kimberly, S. Baughman, R.J. Cabin, J.E. Cohen, N.C. Ellstrand, D.E. McCauley, P. O'Neil, I.M. Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annual Review of Ecological Systems **32**: 305-332.

Stearns, S.C. 1976. Life-history tactics: A review of the ideas. The Quarterly Review of Biology **51**: 3-47.

Tsutsui, N., and A. Suarez. 2002. Reduced genetic variation and the success of an invasive species. Proceedings of the National Academy of Science USA **97**: 5948-5953

U.S. Congress, Office of Technology Assessment, Harmful Non-Indigenous Species in the United States, OTA-F-565 Washington, DC: U.S. Government Printing Office, September 1993.

Vitousek, P.M., C.M. D'Antonio, L.L. Loope, and R. Westbrooks. 1996. Biological invasions as global environmental change. American Scientist **84**: 218–228.

CHAPTER 2: INVASION-ASSOCIATED LIFE HISTORY OF THE LIVEBEARING FISH FAMILY *POECILIIDAE*: HOW DO INVASIVE AND NON-INVASIVE SPECIES DIFFER?[1]

[1] Troendle NJ and Mauricio R. To be submitted to *Ecology Letters*

Abstract

Although invasive species have been the cause of considerable environmental concern for decades, a consensus scientific view as to what biological mechanisms make a species invasive has not yet emerged. One hypothesis is that life history can drive invasive ability, with species with life history traits leading to a high intrinsic rate of increase having an advantage as invasive species. This hypothesis has not been adequately tested, particularly among vertebrate species. Using museum specimens, I compared 11 invasive species from the fish family Poeciliidae to 11 closely related, non-invasive species. An investigation of several life history traits, including embryo size, embryo number, reproductive weight and body size, showed that invasive species have significantly more offspring but that those offspring are smaller than in non-invasive species. This tradeoff could lead to rapid rates of increase in invasive species allowing them to quickly get a foothold and expand when introduced to a new location.

Introduction

Invasive species have been an issue of great concern to ecologists for several decades as they lead to habitat degradation, endemic species decline and, in some cases, extinction (Clavero, 2005; Gurevitch & Padilla, 2004; Mooney & Cleland 2001; Dextrase & Mandrake, 2006). However, despite doing such damage, invasive species do provide opportunities to study evolution as it occurs as invasive species face novel selection pressures and are often introduced widely into various habitats making them natural experiments in evolution (Sakai *et al.* 2001; Huey *et al.* 2005). However, despite the amount of research invested we still do not really know what types of traits make a species a good invader.

Baker (1974) described a set of characteristics that he believed were required for a plant species to be a successful invader. Baker (1974) proposed a set of traits employed by "weed" species, including rapid growth, high fecundity, short generation time, diverse dispersal methods, self-compatibility and a wide range of ecological tolerance. Most of the traits Baker (1974) identified focused on life history traits, traits that affect the species ability to grow and reproduce (Stearns, 1992). Since then, many workers have explored the link between life history and invasion in plants (Sakai *et al.* 2001; Castro-Diez *et al.* 2011; Kolar & Lodge, 2002). However, vertebrate taxa are highly underrepresented in studies of invasion and life history (Olden *et al.* 2006, Vila-Gispert *et al.* 2005).

Fish species make excellent candidates for invasion studies because there are many invasive fish species. Pimentel (2011), for example, reported that of the 138 invasive vertebrates in Brazil, 80% of them were fish species. Of the 100 worst invasive species list produced by the Invasive Species Specialist Group, eight are fish species and fish make up one-third of the vertebrate species on that list (Lowe *et al.* 2000).

There have been several approaches to studying life history and its link to invasiveness. Rosecchi *et al.* (2001) used two species of Cyprinid fishes and attempted to predict invasiveness based upon life history traits measured in these two species. They found that while the two species had some differences in life history traits they had both become invasive. Olden *et al.* (2006) measured life history traits for 90 species of fish from 15 families in the Colorado River basin and were able to classify them into three groups having different life history strategies. However, they found that invasive species did not display a single life history strategy but several. They did however, find that the life history strategies of invasive species differed from those of non-invasive species.

While Olden *et al.* (2006), was unable to find a consistent life history strategy employed by invasive species their study is focused on fish from highly diverse taxa spanning 15 families. It is likely that such diverse taxonomic groups would display different life history strategies. A more focused study investigating a single fish family might be more suitable for examining the life history differences between invasive and non-invasive fish species.

One study that was able to successfully predict invasiveness based on life history traits, was done on the genus *Pinus* by Rejmanek and Richardson (1996). They measured the life history traits of 24 species of pine trees choosing species that were both invasive and non-invasive. They used this data to generate predictors for invasiveness. They were then able to use these predictors to correctly classify the invasive ability of non-pine species as well. One of the strengths of this study is that it looked at multiple species from a single family. As a result, they were able to make comparisons of the life history traits of invasive species and non-invasive species. While this study was effective at showing the link between life history and invasiveness, studies of this type are rare in vertebrate taxa.

The family of live bearing fishes, Poeciliidae provide an excellent system for studying invasion-associated life history. Due to intentional introductions for mosquito control and their popularity in the aquarium trade there are multiple invasive species in the family (Sakai *et al.* 2001). These invasive species can be compared to closely related non-invasive species. Also, because they are live-bearing fishes and females become visibly gravid it is relatively simple to pick out pregnant female fish to be dissected. Also Poeciliidae have been used as a model of life history evolution as it is possible to count and stage each embryo (Reznick & Endler, 1982).

It has been suggested that the two invasive species of Poeciliidae, in the genus *Gambusia*, *G. affinis* and *G. holbrooki* employ different life history strategies than the approximately 50 other species in the genus (Langerhans unpublished work). Langerhans found that these two species had increased fecundity, shorter inter-brood interval, earlier maturity and smaller offspring (unpublished data).

The purpose of this study is to determine the role of life history in invasiveness in a vertebrate taxa, the Poeciliid fishes. I focus on three specific questions in this study. First, are the life history strategies of invasive species different from their non-invasive relatives? Second, are these life history strategies unique to each genus or shared by Poeciliidae? And third, are these differences ones that would benefit an invasive species by increasing the potential rate of increase and spread of the species in a new environment? I hypothesize that changes in life history in invasive species will favor life history strategies that increase the potential rate of increase and spread of the species in a new environment.

Methods

I performed life history dissections on twenty-two species from the family Poeciliidae. I identified and chose eleven invasive species from the Invasive Species Specialist Group, the United States Geological Survey, and Fishbase (Invasive Species Specialist Group, 2016; Nico *et al.* 2016; Froese and Pauly, 2015). Using several phylogenies, I identified appropriate non-invasive species with which to compare the invasive species (Hrbek *et al.*, 2007; Alda *et al.*, 2012; Jones *et al.*, 2013; Lydeard *et al.*, 1995). I obtained fish from University of Michigan Ann Arbor, Academy of Natural Sciences Philadelphia, Tulane University, National Museum of Natural History, Smithsonian Institute, Texas Natural History Science Center, and Florida University Ichthyology Collections. For each of the twenty-two species selected I obtained at least two lots, attempting to choose lots from different parts of the species' range. For each lot chosen, I selected eight to ten gravid females for study (Table S.2.1).

I dissected eight to ten gravid females from each of the lots obtained using the protocol of Reznick and Endler (1982) with a few modifications. Because I used museum specimens I were not able to dry them so all measurements are wet weights. Also because some dehydration occurred while embryos were counted and staged and the degree of desiccation depended on the amount of time it took for the dissection, I rehydrated the embryos and reproductive tissue in deionized water for 30 seconds and then patted dry prior to weighing them. When I found more than one brood at different stages of development each stage was counted and weighed separately. For each fish I measured the total wet weight, length, embryo number, embryo wet weight, total reproductive wet weight, wet weight of reproductive tissue and the number of broods present.

Data Analysis

I transformed my data to better conform to statistical assumptions. In order to determine if there is a difference in the life history strategies of invasive and non-invasive Poeciliid species, I used JMP (SAS, 2012), to perform a Standard Least Squares Means Analysis investigating the effects of invasion status, genus, species, lot, and their interactions with adult weight and stage of development as covariates on the fecundity (measured as the number of embryos), reproductive weight, and embryo size (measured as the wet weight of all embryos divided by fecundity). I used these population least squares means to generate four canonical axes from the invasion status and the genus, invasion status interaction. Finally, I used these canonical axes in a MANCOVA with population, as the unit of replication and invasion status, genus and species were the effects of interest. Effect size, (partial η^2) was calculated using the Wilkes Lambda method (Langerhans *et al.* 2004).

In order to determine if life history was conserved across the four genera studied I used the canonical axes generated from the invasion status, genus and the interaction between these two effects. The canonical axis that I generated from invasion status represents the shared effect of invasion status (Figure 2.1, x-axis). I generated the unique axis of divergence by using the first canonical axis for the interaction between genus and invasion status (Figure 2.2, y-axis). I also used the canonical axes to investigate the correlations between them and the life history traits of interest in order to determine the effects of invasion status and the interaction between invasion status and genus.

As another test of the consistency of life history strategies for invasive species, I performed a discriminant function analysis in three different ways. A discriminant function analysis uses characteristics, in this case, life history traits, to predict group membership, in this

case, invasive or non-invasive species. I used JMP to perform a linear discriminant function analysis using embryo number and embryo weight adjusted for adult fish size to generate discriminant functions using the genus *Gambusia* due to the larger number of lots available for this genus. I then used these functions to predict the invasion status of individual fish across all four genera. I then repeated this analysis using *Poecilia* to generate the predictors. Finally, I performed a linear discriminant function analysis using the leave one out method and the population least squares means to predict the invasion status of each population rather than each individual.

Results

There are several significant differences between the life history strategies of invasive species and non-invasive species. In the twenty-two species examined, species that are considered invasive had significantly higher fecundity (*i.e.,* more embryos) than their non-invasive relatives. The average fecundity for non-invasive species was 21.94 embryos while invasive species had, on average, 34.94 embryos (ANCOVA, $F_{1, 605} = 6.92$, $p = 0.0087$).

This increase in fecundity in the invasive species was balanced by a tradeoff in embryo size. Non-invasive species had significantly higher average embryo weight (8.73 mg), while invasive species offspring averaged 4.62 mg each (ANCOVA, $F_{1, 605} = 54.94$, $p < 0.0001$). I also observed a decrease in total reproductive weight in the invasive species (ANCOVA, $F_{1, 605} = 14.82$, $p = 0.0001$), with invasive species having an average reproductive wet weight of 157.9 mg and non-invasive species averaging 223.7 mg.

There was, however no significant difference in the reproductive allotment (the proportion of body weight made up of reproductive material) between invasive and non-invasive

species (Figure 2.1). Since reproductive allotment did not change but total embryo weight decreased in the invasive species I also tested body size and found that invasive species are significantly smaller than non-invasive species (ANCOVA, $F_{1, 606} = 29.83$, $p < 0.0001$).

I found that species, genus and invasion status, as well as the interaction between genus and invasion status all had significant effects on Poeciliid life history traits (Table 2.1). The largest effect was that of species, but both genus and invasion status and the interaction between genus and invasion status also had significant effects on life history (Table 2.1). When I examine the canonical axes that were generated from the invasion status and genus by invasion status interactions the major difference in life history is characterized by an increase in fecundity in the invasive species balanced by a tradeoff in embryo size. There is a negative correlation between embryo size and embryo number (Table 2.2). A negative correlation also exists between fecundity and embryo size in the first unique canonical axis (Table 2.2) as well as embryo weight and reproductive weight (Table 2.2).

The genera *Gambusia* and *Poecilia* displayed the same tradeoff between fecundity and offspring size, with invasive species having more small offspring than non-invasive species when controlled for adult fish size (Figure 2.1). However, the genera *Xiphophorus* and *Limia* did not show this tradeoff. For these two genera invasive species have significantly fewer offspring and significantly lower reproductive allotment than non-invasive species, but with no change in embryo size.

Overall, invasive species from the genera Gambusia and *Poecilia* had higher fecundity and lower embryo size, similar to the family level results (Figure 2.2).
Non-invasive species have lower fecundity but their offspring are significantly larger than the offspring of invasive species. There was no significant correlation between either fecundity or

embryo weight and reproductive weight. However, the species from the genera *Xiphophorus* and *Limia* fall in the intermediate range of the shared axis of divergence. These two genera do not display this tradeoff in the invasive range. The unique effect is shown by the first canonical axis of the interaction between invasion status and genus (Figure 2.2, y-axis). The first canonical axis for the interaction between genus and invasion status re-emphasized the shared effect pattern of a tradeoff between fecundity and embryo weight, for the genera *Gambusia* and *Poecilia* but showed no such shifts in the genera *Xiphophorous* or *Limia*. For the unique effect there is a negative correlation between embryo weight and reproductive weight (Table 2.2).

The overall trend in body size seems to be driven by the genus *Poecilia*. Invasive species of this genus are much smaller than their non-invasive relatives. However, the other three genera display the opposite trend to different degrees (Figure 2.1). Since body size and reproductive weight are highly correlated, I expect species that grow to larger sizes to have higher reproductive weight. While I do see this pattern in our analyses the genera *Limia* and *Xiphophorus* have higher reproductive allotment in the smaller species than in the larger. So while the larger species have higher reproductive weights because they are larger, it is not so much greater as expected because the smaller species have a higher percentage of their bodies made up of reproductive material.

Discriminant Function Analysis

I was able to correctly predict invasion status based upon life history traits, which indicates that the evolutionary response of invasive species is repeated in multiple species across several populations. This predictability could indicate parallel evolution of life history traits when Poeciliidae become invasive.

Using the discriminant functions generated with *Gambusia* as predictors I were able to correctly predict invasion status 73.3% of the time with increased accuracy within the predictor genus and within invasive species (Table 2.3a). When I used *Poecilia* as the predictor genus the accuracy of prediction increased to 74.7% again with higher fidelity within the predictor genus and within invasive species (Table 2.3b). When I performed the leave one out discriminant function analysis I were able to correctly predict the invasion status of 79.4% of populations again with higher accuracy in the invasive populations than the non-invasive ones (Table 2.3c).

Discussion

Our study found significant effects of invasion status on life history strategies in *Poeciliid* fishes. Across the four genera of livebearers that I sampled, there was a significant shared effect characterized by a trade-off in fecundity and embryo size. Invasive species generally had more offspring than their non-invasive relatives but those offspring were significantly smaller. I did not however, find any significant increase in reproductive allotment in the invasive species. If fecundity increases but reproductive allotment, the amount of resources devoted to reproduction, does not increase as well then something else has to be sacrificed to make up for the increased energy required by producing more offspring (Elnum & Fleming, 2000).

The tradeoff that I observed between fecundity and offspring size is extremely common (Evans *et al.* 2011). It is common for organisms to increase fecundity while sacrificing offspring size despite the fact that increased embryo size leads to decreased juvenile mortality. An organism can produce many small embryos for the same amount of energy that it takes to produce a few large ones. While these smaller offspring may be less fit individually, the overall fitness of the adult may be equal or higher in highly disturbed areas or areas with many predators

where infant mortality will be high regardless of size (Sakai *et al.*, 2001; MacDougall & Turkington, 2005; Didham *et al.* 2005).

While species had a large effect on the life history traits of Poeciliidae our study was focused more on broader taxonomic groups. There are large species-specific differences in life history strategies within the family *Poeciliidae*. For example, the females of the species *Poecilia picta* only grows to be around 5cm in length and produce 10-25 young per brood (Wischnath, 1993; Keith *et al.*, 2000). On the other hand, *Poecilia mexicana* grow to be up to 11cm in length and produce 30-80 young per brood (Greenfield & Tomerson, 1997). However, as this study was aimed at discovering any shared patterns of life history evolution associated with invasion I were more interested in the genus and family level patterns. For this reason, I focused on the effects of invasion status and the interaction between invasion status and genus. These represent their shared and unique effects on life history.

When I look at the genus level, not only do I see the tradeoff between fecundity and embryo size in our shared effect, the unique effect reiterate it. I see that the genera *Poecilia* and *Gambusia* follow this pattern with invasive species being low on the axis and non-invasive species high. However, species from the genera *Xiphophorus* and *Limia* group toward the center regardless of invasion status. Invasive *Xiphophorus* and Limia exhibit very different life history traits in the invasive range than those from *Poecilia* and *Gambusia.*

While this does not support the idea that the family *Poeciliidae* has a common life history response for invasive species, one possible explanation could be the degree of invasiveness of these different genera. For example, *Gambusia affinis* and *holbrooki* are collectively considered to be the most invasive freshwater fish in the world (Pyke, 2008). Likewise, several members of the genus *Poecilia* are extremely invasive, like the guppy *Poecilia reticulata,* which has been

extremely successful as an invader (Lindholm *et al.* 2005; Bambaradeniya, 2002). However, most of the *Xiphophorus* and *Limia* species that are successful invaders tend to be found in only one or two locations, and those tend to be hotspots like Hawaii, and Florida, where there are dozens of invasive species and introduced species are highly successful (USGS, 2016).

Of the estimated 138 non-indigenous fish species introduced to the United States a vast majority of them have become established in regions with mild climates like Florida, Hawaii, and California (Pimentel *et al.*, 2005). These invasion hotspots are susceptible to plant and animal invaders. For example, of the estimated 3,448 plant species in Florida 27% are non-native (Ward, 1990). According to Eldridge and Miller (1997), Hawaii is home to approximately 2690 plant species, 35% of which are invasive. Florida is also home to at least 50 invasive fish species and Hawaii has over 30, (Courtenay, 1997; Maciolek, 1984). Drake and Lodge (2004) proposed that invasion hot spots might be a product of increased ship traffic. It has also been proposed that the presence of invasive species may increase the likelihood of future species invasions (Ricciardi, 2001); this would mean that these hotspots could be easier for new species to invade because they are already homes to other invasive species. So, *Limia* and *Xiphophorus* invasions, which are limited to these types of environments, might be due more to the environments susceptibility than to the species characteristics. Contrastingly, species like *Gambusia affinis*, *Gambusia holbrooki* and *Poecilia reticulate* are extremely widespread and found on multiple continents (Pyke, 2008) tolerate highly diverse environments and may need life history strategies that promote rapid establishment and spread to be successful.

This may indicate that the degree of invasiveness has an effect on life history. For species that are highly invasive and have become widespread like those in the genera *Gambusia* and *Poecilia* these changes to life history make them better at becoming established in new

environments. However, species that are only invasive in one or two locations such as these invasion hotspots may only have become established due to the ecological conditions in those areas. If some metric of invasiveness could be generated, I might observe a correlation between invasiveness and the degree to which the species exhibits invasive life history strategies.

The discriminant function analysis also lends validity to the consistency of these effects. The high rate of predictability indicates that the effect is consistent at least in the genera with more widespread and aggressively invasive species. Using the life history traits (embryo number, reproductive allotment, and embryo size), I were able to correctly predict whether or not a particular population was invasive or not 79.41% of the time with the leave one out method. This level of prediction indicates that these life history traits are strongly correlated with invasion status. The predictability also seems to be higher for the genera that are highly invasive (*Gambusia* and *Poecilia*), with lower levels of correct prediction for less invasive species.

Not only does our discriminant function analysis support our other analyses in showing the general trend of invasive life history, but it also provides a valuable management tool when dealing with potentially invasive species. Using the discriminant functions that I have generated for *Poeciliidae* researchers can determine whether or not a species has high potential for becoming invasive. This will allow managers to identify species of concern. By dissecting a few dozen fish from the species of interest and using the discriminant functions I have generated managers could determine whether or not that species that could become invasive if introduced to a new environment.

I found that whether or not a species is invasive has a significant effect on the life history strategy that species employs. The shared pattern shows invasive species having more offspring of smaller size than their non-invasive relatives. However, this pattern did not hold for all

genera. Fish from the genera *Xiphophorus* and *Limia* did not display this trade-off. I hypothesize that this could potentially be due to the lower intensity of invasion of these groups than the other *Poeciliidae* that were sampled. Further investigation needs to be carried out to determine if there are, perhaps other life history traits that affect these species invasive ability. It could also be that these species are not pre-adapted with these invasive-type life history strategies, but rather their life histories are highly variable and either shift plastically or evolve rapidly when these species actually become invasive. In order to examine this a comparison should be made between invasive *Poeciliidae* species in their native range and their invasive range. Such a study could help to shed light on the differences in how different genera become successful invaders.

References

Baker, H.G. 1974. The evolution of weeds. Annual Review of Ecology and Systematics **5**: 1-24.

Bambaradeniya, C.N. 2002. The status and implications of alien invasive species in Sri Lanka. Zoos' Print Journal **17**: 930-935.

Castro-Díez, P., O. Godoy, A. Saldaña, and D.M. Richardson. 2011. Predicting invasiveness of Australian acacias on the basis of their native climatic affinities, life history traits and human use. Diversity and Distributions **17**: 934-945.

Clavero, M. and E. Garcia-Berthou. 2005. Invasive species are a leading cause of animal extinctions. Trends in Ecology and Evolution **20**: 110-110.

Courtenay, W.R. 1997. Nonindigenous Fishes. Strangers in Paradise. Island Press, Washington, DC.

Dextrase, A.J., and N.E. Mandrak. 2006. Impacts of alien invasive species on freshwater fauna at

risk in Canada. Biological Invasions **8**: 13-24.

Didham, R.K., J.M. Tylianakis, M.A. Hutchison, R.M. Ewers, and N.J. Gemmell. 2005. Are invasive species the drivers of ecological change? Trends in Ecology and Evolution **20**: 470-474.

Drake, J.M., and D.M. Lodge. 2004. Global hot spots of biological invasions: Evaluating options for ballast-water management. Proceedings of the Royal Society of London B Biological Sciences **271**: 575–80.

Einum, S. and I.A. Fleming. 2000. Highly fecund mothers sacrifice offspring survival to maximize fitness. Nature **405**: 565-567.

Eldredge, L.G. and S.E. Miller. 1997. Numbers of Hawaiian species: supplement 2, including a review of freshwater invertebrates. Bishop Museum Occasional Papers **48**: 3–32.

Fernando A., R.G. Reina, I. Doadrio, E. Bermingham. 2013. Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Molecular Phylogenetics and Evolution **66**: 1011-1026.

Greenfield, D.W. and J.E. Thomerson. 1997. Fishes of the Continental Waters of Belize. University Press of Florida, Gainesville, FL.

Gurevitch, J. and D.K. Padilla. 2004. Are invasive species a major cause of extinctions? Trends in Ecology and Evolution **19**: 470-474.

Hrbek, T., J. Seckinger, and A. Meyer. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Molecular Phylogenetics and Evolution **43**: 986-998.

Huey, R.B., G.W. Gilchrist, and A.P. Hendry. 2005. Using invasive species to study evolution. Pp. 138-164 *in*: D.F. Sax, J.J. Stachowicz, and S.D. Gaines (editors). Species Invasions: Insights to Ecology, Evolution and Biogeography. Sinauer Associates, Sunderland, MA.

Jones, J.C., S. Fan, P. Franchini, M. Schartl, and A. Meyer. 2013. The evolutionary history of *Xiphophorus* fish and their sexually selected sword: a genome-wide approach using restriction site-associated DNA sequencing. Molecular Ecology **22**: 2986-3001.

Keith, P., P.Y. Le Bail and P. Planquette. 2000. Atlas des poissons d'eau douce de Guyane. Tome 2. Collection Patrimoines Naturels **43**: 286-286.

Kolar, C.S. and D.M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. Science **298**: 1233-1236.

Langerhans, R.B., C.A. Layman, A. Shokrollahi, and T.J. DeWitt. 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution **58**: 2305-2318.

Lindholm, A.K., F. Breden, H.J. Alexander, W. Chan, S.G. Thakurta, and R. Brooks. 2005. Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. Molecular Ecology **14**: 3671-3682.

Lydeard, C., M.C. Wooten, and A. Meyer. 1995. Molecules, morphology, and area cladograms: a cladistic and biogeographic analysis of *Gambusia* (Teleostei: Poeciliidae). Systematic Biology **44**: 221-236.

MacDougall, A.S. and R. Turkington. 2005. Are invasive species the drivers or passengers of change in degraded ecosystems? Ecology **86**: 42-55.

Maciolek, J.A. 1984. Exotic fishes in Hawaii and other islands of Oceania. Pp. 131-136 *in*: W.R. Courtenay, and J.R. Stauffer Jr. (editors). Distribution, Biology, and Management of

Exotic Fishes. Johns Hopkins University Press, Baltimore, MD.

Mooney, H.A. and E.E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Science, USA **98**: 5446-5451.

Olden, J.D.P., N. LeRoy, K.R. Bestgen. 2006. Life-history strategies predict fish invasions and extirpations in the Colorado River Basin. Ecological Monographs **76**: 25-40.

Pimentel, D., R. Zuniga, and D. Morrison. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics **52**: 273-288.

Pyke, G. H. 2008. Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. Annual Review of Ecology, Evolution, and Systematics **39**: 171-191.

Radford, I.J. and R.D. Cousens. 2000. Invasiveness and comparative life-history traits of exotic and indigenous *Senecio* species in Australia. Oecologia **125**: 531-542.

Reznick, D. and J.A. Endler. 1982. The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). Evolution **36**: 160-177.

Reznick, D.N., F.H. Rodd, and M. Cardenas. 1996. Life-history evolution in guppies (*Poecilia Reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. The American Naturalist **147**: 319-338.

Ricciardi, A. 2001. Facilitative interactions among aquatic invaders: is an 'invasional meltdown' occurring in the Great Lakes? Canadian Journal of Fisheries and Aquatic Sciences **58**: 2513-2525.

Riesch, R., R.A. Martin, and R.B. Langerhans. 2013. Predation's role in life-history evolution of a livebearing fish and a test of the Trexler-DeAngelis model of maternal provisioning. The American Naturalist **181**: 78-93.

Roff, D.A. 2002. Life History Evolution. Sinauer Associates, Sunderland, MA.

Rosecchi, E., F. Thomas, and A.J. Crivelli. 2001. Can life-history traits predict the fate of introduced species? A case study in two cyprinid fish in southern France. Freshwater Biology **46**: 845-853.

Sakai A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, With A. Kimberly, S. Baughman, R.J. Cabin, J.E. Cohen, N.C. Ellstrand, D.E. McCauley, P. O'Neil, I.M. Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annnual Review of Ecological Systems **32**: 305-332.

SAS Institute Inc. 2012. JMP® 10 Modeling and Multivariate Methods. SAS Institute Inc., Cary, NC.

Stearns, S.C. 1976. Life-history tactics: A review of the ideas. The Quarterly Review of Biology **51**: 3-47.

U.S. Geological Survey. 2016. Nonindigenous Aquatic Species Database. Gainesville, FL.

Vila-Gispert, A., C. Alcaraz, and E. Garcia-Berthou. 2005. Life-history traits of invasive fish in small Mediterranean streams. Biological Invasions **7**: 107-116.

Ward, D. B. 1990. How many species are native to Florida? Palmetto **9**: 3-5.

Wischnath, L. 1993. Atlas of Livebearers of the World. T.F.H. Publications Inc., Neptune City, NJ.

Tables and Figures

Table 2.1 Multivariate Analysis of Covariance for Poeciliid Life History Traits

Factor	F	df	P	Partial Variance Explained (%)
Species	2.0628	56, 169	0.0002	63.62
Genus	2.2853	12, 114	0.0121	24.8
Invasion Status	3.2685	4, 43	0.02	23.32
Genus X Invasion Status	2.3334	12, 114	0.0104	25.2

Note: MANCOVA results for life history analysis. F statistics p –values were calculated in JMP.

Partial eta square values were calculated as described in Riesch *et al.* 2013.

Table 2.2 Partial Correlations of Life History Traits

	Fecundity	Reproductive Weight	Average Embryo Weight
Fecundity	1	<0.0001	<0.0001
Reproductive Weight	0.9044	1	0.4407
Average Embryo Weight	-0.9495	0.846	1

Note: values are R^2 correlation coefficients, below the diagonal, generated from Multivariate correlation analysis in JMP. P values are above the diagonal.

Table 2.3 Discriminant Function Analyses

Prediction Method	Group Predicted	% Correctly Classified
A. Gambusia	Gambusia	79.89
	Poecilia	67.4
	Limia	63.49
	Xiphophorus	64.33
	Invasive	77.89
	Non-invasive	61.88
	Total	73.3
B. Poecilia	Gambusia	82.59
	Poecilia	68.03
	Limia	60.32
	Xiphophorus	64.33
	Invasive	80.15
	Non-invasive	61.25
	Total	74.73
C. Leave One Out	Gambusia	89.19
	Limia	50
	Poecilia	72.22
	Xiphophorus	66.67
	Invasive	88.89
	Non-invasive	68.75
	Total	79.41

Note: The table shows the fidelity of three different discriminant function analyses. Each section shows the prediction rates of each genus, invasive species, non-invasive species and all predictions for each prediction method. A. summarizes predictions when the genus *Gambusia* is used as to make predictions. B. uses the genus *Poecilia* to make predictions and C. uses the leave one out method. Predictions are shown as the percentage of fish correctly classified for sections A and B. and percentage of populations correctly classified for C.

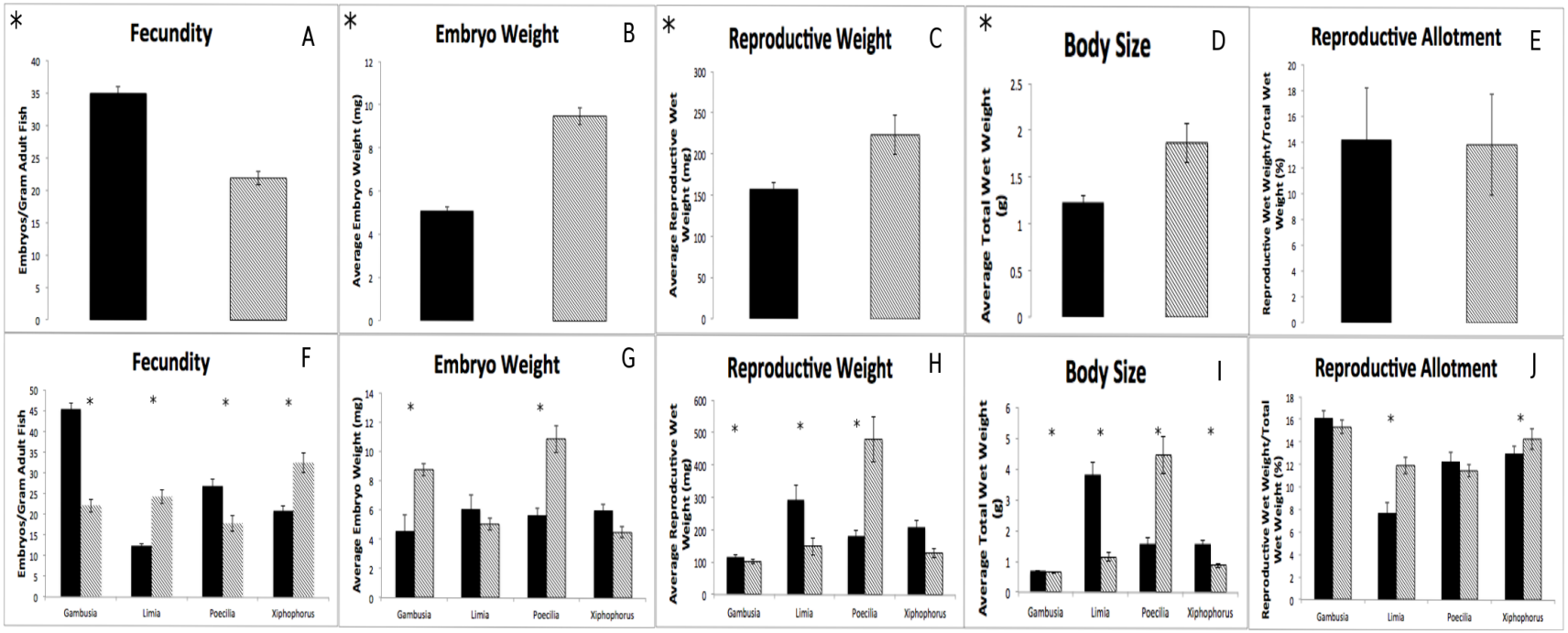

Fecundity
A
Embryos/Gram Adult Fish
Embryo Weight
B
Average Embryo Weight (mg)
Reproductive Weight
C
Average Reproductive Wet Weight (mg)
Body Size
D
Average Total Wet Weight (g)
Reproductive Allotment
E
Reproductive Wet Weight/Total Wet Weight (%)
Fecundity
F
Embryos/Gram Adult Fish
Embryo Weight
G
Average Embryo Weight (mg)
Reproductive Weight
H
Average Reprodcutive Wet Weight (mg)
Body Size
I
Average Total Wet Weight (g)
Reproductive Allotment
J
Reproductive Wet Weight/Total Wet Weight (%)
Gambusia
Limia
Poecilia
Xiphophorus

Figure 2.1. Invasive species are graphed in black non-invasive species are striped. * denotes a statistically significant difference. For graphs with only two bars an * in the upper corner indicates the difference is significant. Panel A shows the average number of embryos controlled for the size of the adult fish. Panel B shows reproductive allotment Reproductive weight/total weight of the adult female. Panel C shows the average embryo weight/total weight of the adult female.

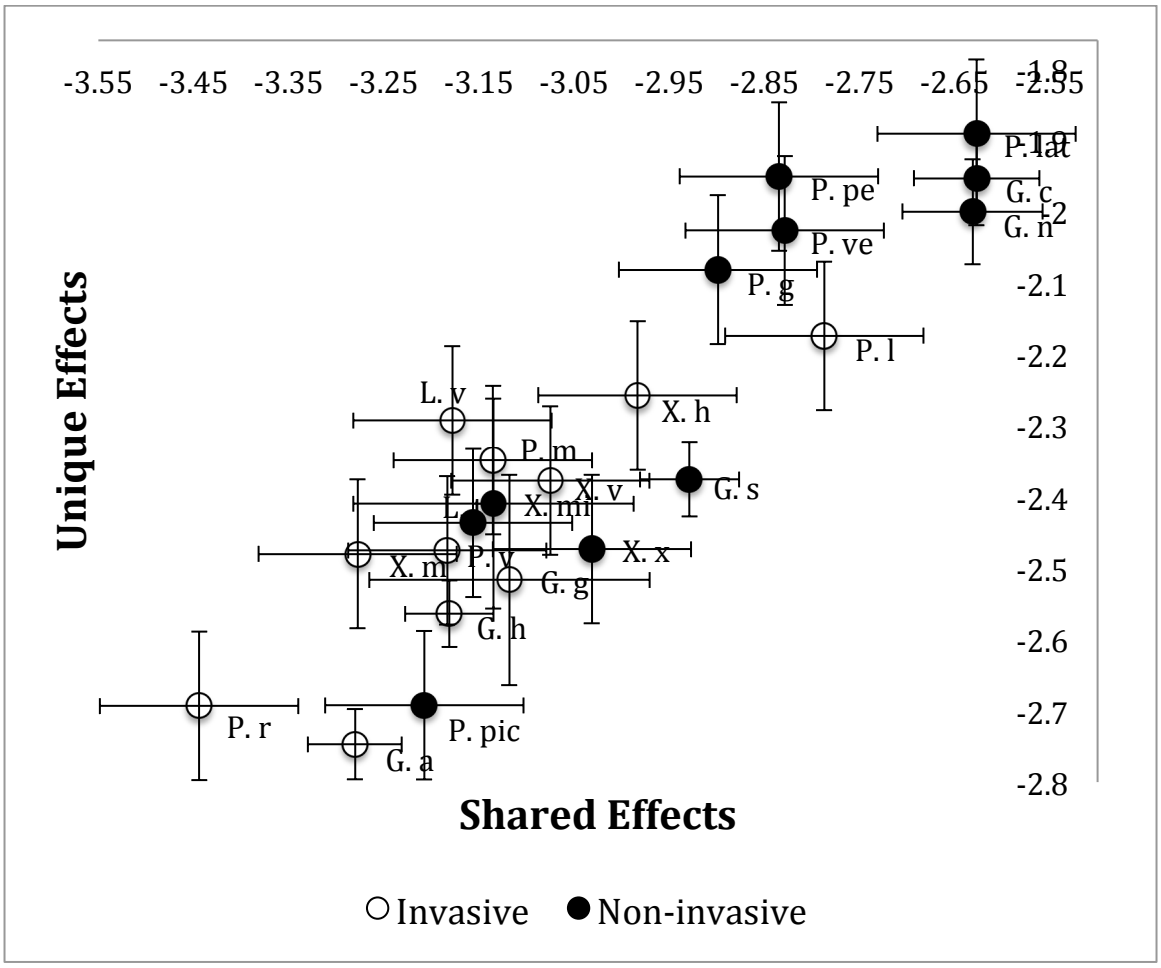


Figure 2.2: Multivariate variation in Poeciliidae life history traits, shared effects (from invasion status canonical axis), unique effects (from interaction of invasion status and genus first canonical axis). Means and standard errors depicted. Closed circles are non-invasive species and open circles are invasive species. The X-axis depicts the shared effect as it increases embryo number decreases, reproductive weight increases and average embryo weight increases. The Y-axis shows the unique effect. As it increases fecundity decreases, reproductive weight decreases, and average embryo weight decreases. Species: Gambusia affinis (G. a), Gambusia holbrooki (G. h), Gambusia geiseri (G. g), Gambusia clarkhubbsi (G. c), Gambusia nobilis (G. n), Gambusia speciosa (G. s), Poecilia reticulate (P. r), Poecilia Mexicana (P. m), Poecilia vivipara (P. v), Poecilia latipinna (P. l), Poecilia picta (P. pic), Poecilia gilli (P. g), Poecilia velifera (P.

ve), Poecilia petenensis (P. pe), Poecilia latipunctata (P. lat), Limia vittata (L. v), Limia tridens, (L. t), Xiphophorus helleri (X. h), Xiphophorus maculatus (X. m), Xiphophorus variatus (X. v), Xiphophorus milleri (X. mi), and Xiphophorus xiphidium (X. x).

Appendix A

Table S.2.1: Species used for Life history analysis

Genus	Species	Invasion Status	Lots Dissected
Gambusia	affinis	Invasive	9
Gambusia	holbrooki	Invasive	10
Gambusia	geiseri	Invasive	1
Limia	vittata	Invasive	2
Poecilia	latipinna	Invasive	2
Poecilia	mexicana	Invasive	2
Poecilia	reticulata	Invasive	2
Poecilia	vivipara	Invasive	2
Xiphophorus	helleri	Invasive	2
Xiphophorus	maculatus	Invasive	2
Xiphophorus	variatus	Invasive	2
Gambusia	clarkhubbsi	Non-invasive	5
Gambusia	nobilis	Non-invasive	4
Gambusia	speciosa	Non-invasive	8
Limia	tridens	Non-invasive	2
Poecilia	gilli	Non-invasive	2
Poecilia	latipunctata	Non-invasive	2
Poecilia	petenensis	Non-invasive	2
Poecilia	picta	Non-invasive	2
Poecilia	velifera	Non-invasive	2
Xiphophorus	milleri	Non-invasive	1
Xiphophorus	xiphidium	Non-invasive	2

Note: Shows the genus and species for each population dissected for this project as well as their invasion status and how many populations were dissected for each species.

Table S.2.2: Descriptive statistics of Invasive and Non-invasive Poeciliid Life History

Invasion Status, Species	N	SL (mm)	TWW (g)	F	RA (%)	EW (mg)
Invasive Species						
G. affinis	83	34.77±3.94	.78±.29	44.17±21.43	17.68±8.50	2.89±1.38
G. geiseri	10	35.64±3.63	.79±.25	27.6±9.79	17.24±5.24	4.76±1.82
G. holbrooki	86	31.49±5.38	.58±.32	19.72±14.02	14.55±6.83	4.02±1.79
Gambusia	179	33.26±4.94	.68±.32	31.49±5.38	16.18±7.69	4.58±14.06
L. vittata	18	53.83±7.81	3.84±1.162	46.28±18.84	7.71±4.09	6.10±3.89
Limia	18	53.83±7.81	3.84±1.62	46.28±18.84	7.71±4.09	6.10±3.89
P. latipinna	20	37.35±2.78	1.27±.30	26.25±12.86	20.67±7.75	10.14±3.48
P. mexicana	16	56.33±11.24	4.35±2.19	54.13±41.66	8.31±3.09	7.08±4.24
P. reticulata	26	30.17±3.90	.63±.24	23.88±11.40	8.95±3.83	2.20±.73
P. vivipara	17	32.55±6.51	.86±.49	24.65±17.97	11.00±5.04	4.21±2.48
Poecilia	79	37.75±11.67	1.59±1.74	30.72±24.97	12.28±7.18	5.65±4.26
X. helleri	21	53.09±6.20	2.94±.95	46.48±21.58	12.07±5.41	7.33±3.64
X. maculatus	16	27.28±2.87	.55±.27	9.5±4.62	9.90±4.62	4.32±2.59
X. variatus	26	36.03±5.54	1.08±.46	31.35±20.01	15.64±4.23	5.99±2.33
Xiphophorus	63	39.49±11.53	1.57±1.18	31.17±22.98	12.99±5.26	6.01±3.07
Total/Avg	339	36.56±9.76	1.23±1.31	32.05±22.55	14.20±7.38	4.62±3.08
Non-Invasive Species						
G. clarkhubbsi	47	32.20±4.85	.60±.38	5.66±4.31	9.45±4.85	9.78±4.85
G. nobilis	40	32.68±4.24	.76±.30	9.93±4.32	16.94±5.05	12.87±4.18
G. speciosa	67	30.15±4.36	.59±.27	20.99±15.10	18.48±8.25	5.57±2.87
Gambusia	154	31.43±4.6	.64±.32	13.44±12.47	15.32±7.48	8.75±4.93
L. tridens	20	35.25±5.02	1.16±.64	28.2±15.34	11.92±3.26	5.04±1.75
Limia	20	35.25±5.02	1.16±.64	28.2±15.34	11.92±3.26	5.04±1.75
P. gilli	13	52.9±10.38	3.83±2.08	39.46±22.61	9.45±2.49	8.85±2.81
P. latipunctata	20	37.72±3.69	1.19±.38	10±2.68	13.10±4.20	14.98±7.85
P. petenensis	17	76.47±18.22	12.37±8.32	69.59±28.90	9.22±4.07	15.08±10.06
P. picta	19	20.54±2.12	.15±.06	8±4.67	11.56±4.60	2.02±.71
P. velifera	20	58.98±8.20	5.51±2.47	50±28.33	13.13±7.54	12.82±8.08
Poecilia	89	48.45±21.62	4.46±5.78	34.25±31.38	11.5±5.2	10.85±8.52
X. milleri	9	33.16±2.34	.87±.19	19.56±6.27	16.00±5.64	6.57±1.16
X. xiphidium	19	32.34±5.93	.89±.47	30.63±12.62	13.47±4.15	3.56±1.52
Xiphophrous	28	32.6±5.02	.89±.39	27.07±12.07	14.29±4.72	4.53±1.99
Total/Avg	291	37.01±14.69	1.86±3.63	22.13±22.36	13.82±6.61	8.73±6.28

Note: Means and standard deviations are shown for female life history traits. Number of individuals dissected (N), Standard Length of the adult in mm (SL), Total wet weight of the adult in g (TWW), Fecundity (F), Reproductive allotment, Reproductive weight/total wet weight given as a % (RA), Average embryo weight in mg (EW).

CHAPTER 3: INVASION-ASSOCIATED LIFE HISTORY SHIFTS IN THE INVASIVE RANGE OF INVASIVE POECILIIDAE[2]

[2] Troendle NJ and Mauricio R. To be submitted to *Evolution*

Abstract

Invasive species have been one of the topics of greatest ecological concern for decades, however our understanding of what types of traits make a species invasive is still incomplete. One theory is that life history can drive invasive ability. Species with life history traits leading to a high intrinsic rate of increase would have an advantage as invasive species. This has been only poorly tested and most studies have focused on plant species, with vertebrate species being highly underrepresented. Also, few studies have been done that examine the shifts in life history that take place in the invasive range of species. I examined life history traits of eleven invasive species comparing their native and invasive ranges in order to determine if life history traits shifted as a result of becoming invasive. My investigation showed that in their invasive range invasive species displayed life history strategies that increased their population growth rate. Body size, fecundity, and reproductive allotment all increased in the invasive range without sacrificing offspring size. This suggests that when a species colonizes a new environment there are shifts in life history that help them to become established more quickly.

Introduction

Invasive species are an issue of great concern and have been found to be one of the leading causes of endemic decline and extinction (Clavero and Garcia-Berthou, 2005; Gurevitch & Padilla, 2004). Invasive species are important drivers of ecological and environmental change (Sakai *et al.* 2001; Huey *et al.* 2005). However, the act of invasion changes not only the native species, but also the invading species itself. An invasive species is subjected to novel environmental and evolutionary pressures making invasive species natural experiments in evolution (Lee, 2002). By observing changes in the invasive range of a species we can investigate evolution in a natural system in real time. Invasive species often seem to evolve and adapt rapidly (Prentis *et al.* 2008). One group of traits that generally evolve rapidly are life history traits (Dlugosch and Parker, 2008; Hariston and Walton, 1986).

Life history traits are the demographic measurements of an organism and figure directly into its fitness (Stearns, 1992). Traits like age at sexual maturity, number and size of offspring, reproductive duration and senescence all directly affect an organism's fitness and the intrinsic rates of increase of a population (Stearns, 1976; Reznick *et al.* 1996). Life history traits are generally constrained by trade-offs either due to limited energy or time (Roff, 2002). For example, given a finite number of resources, an organism must "decide" how to divide up those resources between growth and reproduction. This leads to different suites of life history traits in different environments as different strategies optimize fitness under different ecological conditions. One of the most common life history trade-off is the link between offspring number and offspring size (Lack, 1947; Stearns, 1992). If the amount of resources devoted to reproduction does not increase but clutch size does, then the size of the offspring has to be reduced in order to account for the increase in number (Roff, 2002). One example of this is the

study by Walker *et al.* (2008), which found that in primates, including humans, have decreased birth weight with increased litter size. This trade-off may be important for invasive species, as disturbed habitats are more likely to be colonized by invasive species (Sakai *et al.*, 2001; MacDougall & Turkington, 2005; Didham *et al.* 2005). Unpredictable habitats often lead to bet-hedging strategies to ensure reproduction (Olofsson *et al.* 2009). While it has been proposed that in such unstable environments producing a smaller number of high quality offspring is the better strategy this is not always the case (Morrongiello *et al.* 2012). They argue that if the source of mortality is not dependent on the size of the offspring then increasing fecundity is at the cost of offspring size is the optimal strategy.

The family of live bearing fishes, Poeciliidae, provides an excellent system for studying invasion-associated life history. Due to intentional introductions for mosquito control and their popularity in the aquarium trade there are multiple invasive species in the family (Sakai *et al.* 2001). These invasive species provide an opportunity to study how life history evolves when a species becomes invasive. Also, because they are live-bearing fishes and females become visibly gravid it is relatively simple to pick out pregnant female fish to be dissected. Also, Poeciliidae have been used as a model of life history evolution as it is possible to count and stage each embryo (Reznick & Endler, 1982).

In chapter 2, I investigated the differences in life history between invasive and non-invasive Poeciliidae. I looked at 11 invasive species and 11 non-invasive species from the family and found that invasive species displayed life history strategies that are indicative of high rates of increase and therefore, better chance of survival in dynamic and unstable environments which are inherently more susceptible to invasion (Didham *et al.* 2005). These included

increased fecundity for most invasive species but this was balanced by a reduction in the size of their offspring, which is a common life history tradeoff.

While that study demonstrated that there is an inherent difference between the life history traits of invasive species and those of their non-invasive relatives, it does not address the question of whether or not invasive species change life history strategies when they become invasive. That is, do invasive species alter their life history traits, either through rapid evolution or phenotypic plasticity depending on whether they are located in their native or invasive range? In chapter 2, I investigated the differences in life history between eleven invasive and eleven non-invasive species. However, in order to determine how species change when they become invasive I needed to compare life history in the native and invasive ranges of those eleven invasive species. As the previous study only measured life history in the native range of these species it was unable to address this question.

The primary goal of this study is to determine if invasive Poeciliidae display life history traits that would lead to higher rates of increase in their invasive range than they do in their native range. If this were the case I would expect fecundity and reproductive allotment to be higher in the invasive range than in the native range. As an increase in fecundity is often balanced by a decrease in offspring size, (Evans *et al.* 2011), I would also expect that average offspring size in the invasive range would be smaller than in the native range. However, if this were not the case and life history traits like fecundity and reproductive allotment are already performing at maximum output in invasive species, I would expect to see no differences in these traits between the two ranges. I was also interested to know if all four genera represented in my study adapt to becoming invasive using the same life history strategies or if each genus has different strategies when it becomes invasive.

Methods

I performed life history dissections on eleven species from the family Poeciliidae. I identified and chose eleven invasive species from the Invasive Species Specialist Group, the United States Geological Survey, and Fishbase (Invasive Species Specialist Group, 2016; Nico *et al.* 2016; Froese and Pauly, 2015). I obtained fish from University of Michigan Ann Arbor, Academy of Natural Sciences Philadelphia, Tulane University, National Museum of Natural History, Smithsonian Institute, Texas Natural History Science Center, and Florida University Ichthyology Collections. For each of the eleven species selected I obtained at least two lots from their documented invasive range and two lots from their native range, attempting to choose lots from different parts of each range. For each lot chosen eight to ten gravid females were randomly selected for study (Table S.1).

I dissected eight to ten gravid females from each of the lots obtained using the protocol of Reznick and Endler (1982) with a few modifications. Because museum specimens were used I was not able to dry them so all measurements in this study are of wet weights rather than dry or lean weights. Also because some dehydration occurred while embryos were counted and staged and the degree of desiccation depended on the amount of time it took for the dissection, I rehydrated the embryos and reproductive tissue in deionized water for 30 seconds prior to weighing them. When more than one brood at different stages of development were found, each stage was counted and weighed separately. I collected data for life history traits including standard length (mm), wet weight (g), fecundity, embryo wet weight (g), stage of development of the offspring, reproductive allotment (mass of offspring and reproductive tissue/total mass), and average embryo weight (g), weight of all embryos/number of embryos present).

I transformed my data to conform to statistical assumptions. I then used JMP (SAS, 2012), to perform a MANCOVA investigating the effects of range, genus, species, lot, and their interactions with adult weight and stage of development as covariates on the fecundity, reproductive weight, and embryo size. The whole model being significant, I performed subsequent ANCOVAs and standard least squares means for each of these traits. Finally, I made range contrasts for each genus comparing the invasive range to the native range for each of the three traits using the contrast function in JMP.

In order to test for allometric relationships between body size and fecundity, reproductive weight and embryo size, I analyzed correlations between body size and these three traits. In the initial MANCOVA, body size was included as a covariate in order to remove the variance due to body size differences between species and between individuals of a particular species. However, because the size of reproductive individuals could also be influenced by invasion a separate MANCOVA was performed including total wet weight as one of the traits of interest rather than a covariate. Due to the established positive correlation between reproductive weight and body size, I used an alternate measure for reproductive weight, reproductive allotment, which is the weight of all reproductive material divided by the total body weight in analyses where body weight was a trait of interest rather than a covariate. I calculated effect sizes for both MANCOVAs using Wilks's partial η^2 (Langerhans *et al.* 2004). I generated partial correlations from the MANCOVA for fecundity, reproductive weight, total wet weight, and average embryo weight in order to investigate the relationships between these four traits of interest.

Results

Family Level Results

I found that there were several differences in the life history traits displayed by Poeciliidae in their native and invasive ranges. I found that species, genus, and range all had significant effects on female life history (Table 1). There were significant differences in life history between populations of the same species (MANOVA, $F_{212, 2438} = 8.58$, $p < 0.0001$). There were also significant differences between species and genera (MANOVA, $F_{28, 2201} = 35.18$, $p < 0.0001$ and $F_{12, 1614} = 70.21$, $p < 0.0001$). The range also had a statistically significant effect on life history and there were significant differences in life history traits between the two ranges (MANOVA, $F_{4, 610} = 7.52$, $p < 0.0001$). However, since I am interested in broad scale effects, the effects of range and genus are the focus of this study.

I found that fish in the invasive range had higher fecundity than fish of the same species, collected in their native range (MANCOVA, $F_{1, 613} = 22.08$, $p < 0.0001$). Reproductive weight and reproductive allotment that is the proportion of the total body weight made up of reproductive materials were also significantly higher in the invasive range (MANCOVA, $F_{1, 613} = 14.03$, $p = 0.0051$ and $F_{1, 613} = 9.14$, $p = 0.0026$, respectively). Body weight was significantly higher in the invasive range (ANCOVA, $F_{1, 613} = 6.17$, $p = 0.0132$). Average embryo weight was not significantly different between the two ranges.

My analysis of correlation between the traits of interest revealed that all four of the traits displayed strong correlations with one another. Fecundity was positively correlated with both reproductive allotment ($R^2 = 0.9238$, $p < 0.0001$; Table s.3.3.A) and total body weight ($R^2 = 0.9502$, $p < 0.0001$; Table S.3.3.A). Reproductive allotment was positively correlated with embryo weight ($R^2 = 0.9258$, $p < 0.001$). Reproductive weight was positively correlated with

fecundity ($R^2 = 0.9238$, $p < 0.0001$; Table S.3.3.B), embryo weight ($R^2 = 0.9258$, $p < 0.001$; Table S.3.3.B) and body size ($R^2 = 0.3192$, $p < 0.0001$; Table S.3.3.B). Finally, embryo size was negatively correlated with body size ($R^2 = –0.1223$, $p < 0.001$; Table S.3.3.A).

Genus Level Results

I found that each of the four genera of invasive species from the family Poeciliidae displayed variable life history responses to invasion. For the genera *Limia* and *Xiphophorus,* life history differed depending on range (MANOVA, $F_{4, 31} = 5.58$, $p = 0.0017$ and $F_{4, 101} = 5.15$, $p = 0.0008$). In these two genera range was responsible for 41.87% and 16.94% of the partial variance respectively (Table 3.1). However, for the genera *Gambusia* and *Poecilia,* range had a smaller effect with partial variances of 6.83% and 8.76% respectively (Table 3.1); there was still a significant difference in life history between the ranges (MANOVA, $F_{4, 315} = 5.77$, $p = 0.0002$ and $F_{4, 151} = 3.62$, $p = 0.0075$ respectively).

Range had a significant effect on fecundity of invasive species, though the magnitude and direction varied by genera. In the genera *Gambusia, Limia*, and *Xiphophorus,* I observed a significant increase in fecundity in the invasive range though to varying degrees (ANCOVA, $F_{1, 613} = 7.62$, $p = 0.0059$; $F_{1, 613} = 9.45$, $p = 0.0022$; and $F_{1, 613} = 26.83$, $p < 0.0001$ respectively, Table 3.2). *Poecilia,* however, displayed significantly lower fecundity in their invasive range (ANCOVA, $F_{1, 613} = 4.91$, $p = 0.027$, Table 3.2).

The genera *Gambusia* and *Limia* had significantly higher reproductive allotment in the invasive range than in their non-invasive ranges (ANCOVA, $F_{1, 613}=4.57$, $p = 0.0329$; and $F_{1, 613} = 11.38$, $p = 0.0008$ respectively). However, the genera *Poecilia* and *Xiphophorus* did not display

a significant difference in reproductive allotment between their ranges (ANCOVA, $F_{1, 613} = 2.03$, $p = 0.1545$; and $F_{1, 613} = 0.20$, $p = 0.6512$ respectively).

Only the genus *Gambusia* displayed a significant difference in embryo size with invasive fish having larger offspring than non-invasive fish (ANCOVA, $F_{1, 613} = 9.53$, $p = 0.0021$). *Limia* and *Poecilia* also displayed slight increases in embryo size in the invasive range but the differences were not statistically significant. *Xiphophorus* embryo size tended to be smaller in the invasive range but again this difference was not statistically significant.

The genera *Gambusia* and *Xiphophorus* were significantly larger in their invasive range (ANCOVA, $F_{1, 613} = 14.07$, $p = 0.0002$; and $F_{1, 613} = 22.04$, $p < 0.0001$ respectively). *Gambusia* were on average over 35% larger in the invasive range than they were in the native range while *Xiphophorus* increased by 3.8% in their invasive range. *Poecilia* displayed the opposite effect with fish being 17.8% larger in the native range than the invasive range (ANCOVA, $F_{1, 613} = 5.94$, $p = 0.0151$). *Limia* displayed no significant change in body size between its ranges (ANCOVA, $F_{1, 613} = 0.05$, $p = 0.8305$).

Discussion

Since Baker (1974), proposed the concept of an ideal weed, scientists have explored the link between invasiveness and life history. Baker proposed a set of life history strategies employed by "weed" species characterized by rapid growth, high fecundity, short generation time, diverse dispersal methods, and self-compatibility. While no invasive species meets all of Baker's criteria most display either a subset or some variation on them. Because of this link it has been speculated that life history traits might be used as predictors of invasiveness (Sakai *et al.* 2001; Castro-Diez *et al.* 2011; Kolar & Lodge, 2002).

In chapter 1, I showed that invasive Poeciliidae tend to have life history traits that would seem to help them become established and increase their population size over that of their non-invasive relatives. They found that fecundity was typically higher in invasive species but was balanced by a decrease in embryo size. It is interesting though to ask what changes when a species becomes invasive. Chapter 2, looked specifically at differences between invasive and non-invasive species but did not investigate how invasive species change when they become invasive. In order to do this, comparisons need to be made between the invasive and native ranges of invasive species.

When looking at life history traits in the family Poeciliidae, I found that overall there was an increase in overall fitness and reproduction. Fecundity in the invasive range increased significantly over the native range. This increase in fecundity was not balanced by a tradeoff in embryo size, as there was no significant difference in embryo size but rather by increases in body size and reproductive weight. As there is a well-established positive correlation between body size, and reproductive weight and body size and fecundity (Stearns, 1992), the increase in body size would seem to explain the increase in these reproductive traits. However, I also found that reproductive allotment was significantly higher in the invasive range. This indicates that overall reproduction is increasing more than is explained by allometry. That is, not only are they reproducing more because they are bigger, but they are reproducing more when body size is controlled for.

Typically, growth and reproduction are involved in a tradeoff constrained by the availability of resources. With a given amount of food, so much energy can be produced and that energy must be allocated among several processes, including growth and reproduction. If an organism allocates more resources to growth, it has fewer resources to devote to reproduction. I

do not see that here, however, which indicates that in the invasive range these species are actually doing better as a whole. They either have fewer predators and parasites, which allows them to devote the resources that would have gone to fighting off diseases or escaping predators to growth and reproduction; or they have less competition, which allows them to obtain more food. This allows reproduction and growth to both increase as they have more energy overall and are not forced to trade off between the two.

In my first chapter I also found that for the genera *Gambusia* and *Poecilia*, invasive species displayed a common life history tradeoff. Invasive species had higher fecundity than their non-invasive relatives but this was balanced by a decrease in the size of their offspring. However, for the genera *Limia* and *Xiphophorus*, they found that fecundity was actually lower in the invasive species and their reproductive allotment was reduced as well with no significant change in embryo size. This was considered odd since these species were not displaying any of the life history strategies expected to characterize an invasive.

In my study, however, I see the increases in fecundity in *Limia* and *Xiphophorus* that would be predicted in an invasive species. *Limia* and *Xiphophorus* both had significantly more offspring in their invasive ranges than in their native ranges. This increase in fecundity was paid for in several ways. For *Limia* it corresponded with an increase in reproductive allotment. This indicates that when *Limia* become invasive they devote a greater proportion of their resources to reproduction, producing a greater number of offspring but not sacrificing embryo size, as there was no significant difference in embryo size between the two ranges. *Xiphophorus* however, does not display significantly increased reproductive allotment. Rather the invasive individuals are significantly larger than their non-invasive counterparts. This allows *Xiphophorus* to

produce a greater number of offspring without increasing reproductive allotment or decreasing embryo size.

The genus *Poecilia* yielded some interesting results. They did not display any significant difference in embryo size or reproductive allotment between their two ranges. However, they had significantly lower fecundity in the invasive range. This corresponded with a decrease in body size in the invasive range. While this seems to directly contradict my expectations, it is likely that this is due to a sampling issue with the species *Poecilia mexicana*. It was extremely difficult to find pregnant females from the invasive range of *P. mexicana*. Most populations obtained from museums either had no pregnant females or only a couple per lot. Those females that were pregnant in the invasive range tended to be much smaller than their native range counterparts and have far lower fecundity. These females' embryos were on average two stages earlier than their native counterparts. Coupled with smaller size and decreased fecundity this could indicate that the females sampled in the invasive populations were younger than those in the native populations.

My previous chapter showed that invasive species in the genus *Gambusia* had a significant increase in fecundity when compared to their non-invasive relatives, and that this was balanced by a decrease in the size of their offspring. In my investigation of the invasive and native range I determined that invasive *Gambusia* become even more reproductively successful in the introduced range. I found that there was a significant increase in body size. Since there is a well-established correlation between body size and reproduction it is not surprising that reproductive weight is higher in the invasive range as well. Fecundity is also higher as expected. However, *Gambusia* in their invasive range also displays an increase in reproductive allotment, the proportion of resources dedicated to reproduction. *Gambusia* was also the only genus to

show a significant difference in embryo size between the native and invasive range. However, rather than showing tradeoff with fecundity, embryo size increased in the invasive range.

My study demonstrates the importance of invasive species as natural evolutionary systems. I observed significant changes in the life history strategies of species known to become invasive in their invasive range. This indicates that not only are do many of these fish have life history traits that make them potentially invasive as I showed in my previous chapter, but they are also able to change their life history strategies when they become invasive to make them better suited to their new environment. This could be due to rapid evolution or phenotypic plasticity and further research is necessary to determine which of these processes is responsible for these rapid shifts in life history.

References

Baker, H.G. 1974. The evolution of weeds. Annual Review of Ecology and Systematics **5**: 1-24.

Castro-Díez, P., O. Godoy, A. Saldaña, and D.M. Richardson. 2011. Predicting invasiveness of Australian acacias on the basis of their native climatic affinities, life history traits and human use. Diversity and Distributions **17**: 934-945.

Clavero, M. and E. Garcia-Berthou. 2005. Invasive species are a leading cause of animal extinctions. Trends in Ecology and Evolution **20**: 110-110.

Dextrase, A.J., and N.E. Mandrak. 2006. Impacts of alien invasive species on freshwater fauna at risk in Canada. Biological Invasions **8**: 13-24.

Didham, R.K., J.M. Tylianakis, M.A. Hutchison, R.M. Ewers, and N.J. Gemmell. 2005. Are invasive species the drivers of ecological change? Trends in Ecology and Evolution **20**: 470-474.

Dlugosch, K.M., and I.M. Parker. 2008. Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. Ecology Letters **11**: 701-709.

Evans, J.P., A. Pilastro, and I. Schlupp. 2011. Ecology and Evolution of Poeciliid Fishes. University of Chicago Press, Chicago, IL

Froese, R. and D. Pauly. 2016. FishBase. World Wide Web electronic publication.

Gurevitch, J. and D. K. Padilla. 2004. Are invasive species a major cause of extinctions? Trends in Ecology and Evolution **19**: 470-474.

Hairston, N.G. Jr., and W.E. Walton. 1986. Rapid evolution of a life history trait. Proceedings of the National Academy of Science, USA **83**: 4831-4833.

Huey, R.B., G.W. Gilchrist, and A.P. Hendry. 2005. Using invasive species to study evolution. Pp. 138-164 *in:* D.F. Sax, J.J. Stachowicz, and S.D. Gaines (editors). Species Invasions: Insights to Ecology, Evolution and Biogeography. Sinauer Associates, Sunderland, MA.

Kolar, C. S. and D. M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. Science **298**: 1233-1236.

Lack, D. 1947. The significance of clutch size. 1. Intraspecific variation. Ibis. **89**: 302–352.

Lee, C.E. 2002. Evolutionary genetics of invasive species. Trends in Ecology and Evolution. **17**: 386-391.

Langerhans, R.B., C.A. Layman, A. Shokrollahi, and T.J. DeWitt. 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution **58**: 2305-2318.

MacDougall, A.S. and R. Turkington. 2005. Are invasive species the drivers or passengers of change in degraded ecosystems? Ecology **86**: 42-55.

Mooney, H.A. and E.E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Science, USA **98**: 5446-5451.

Morrongiello, J.R., N.R. Bond, D.A. Crook, and B.B.M. Wong. 2012. Spatial variation in egg size and egg number reflects trade-offs and bet-hedging in a freshwater fish. Journal of Animal Ecology **81**: 806-817.

Nico, L., P. Fuller, G. Jacobs, M. Cannister, J. Larson, A. Fusaro, T.H. Makled and M. Neilson. 2016. *Gambusia affinis*. USGS Nonindigenous Aquatic Species Database, Gainesville, FL.

Olofsson, H., J. Ripa, and N. Jonzen. 2009. Bet-hedging as an evolutionary game: the trade-off between egg size and number. Proceedings of the Royal Society B. **276**: 2963-2970.

Prentis, P.J., J.R.U. Wilson, E.E. Dormontt, D.M. Richardson, and A.J. Lowe. 2008. Adaptive evolution in invasive species. Trends in Plant Science **13**: 288-294.

Reznick, D. and J. A. Endler. 1982. The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). Evolution **36**: 160-177.

Reznick, D.N., F.H. Rodd, and M. Cardenas. 1996. Life-history evolution in guppies (*Poecilia Reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. The American Naturalist **147**: 319-338.

Riesch, R., R.A. Martin, and R.B. Langerhans. 2013. Predation's role in life-history evolution of a livebearing fish and a test of the Trexler-DeAngelis model of maternal provisioning. The American Naturalist **181**: 78-93.

Roff, D.A. 2002. Life History Evolution. Sinauer Associates, Sunderland, MA

Sakai A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, With A. Kimberly, S. Baughman, R.J. Cabin, J.E. Cohen, N.C. Ellstrand, D.E. McCauley, P. O'Neil, I.M.

Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annual Review of Ecological Systems **32**: 305-332.

SAS Institute Inc. 2012. JMP® 10 Modeling and Multivariate Methods. SAS Institute Inc., Cary, NC.

Stearns, S.C. 1976. Life-history tactics: A review of the ideas. The Quarterly Review of Biology **51**: 3-47.

Stearns, S.C. 1992. The Evolution of Life Histories. Oxford University Press, New York NY.

Walker, R.S., M. Gurven, O. Burger, and M.J. Hamilton. 2008. The trade-off between number and size of offspring in humans and other primates. Proceedings of the Royal Society of London B: Biological Sciences **275**: 827-833.

Tables

Table 3.1: Multivariate Analysis of Covariance for Poeciliid Life History Traits

Factor	F	df	P	Partial Variance Explained (%)
A. Gambusia				
Population	6.79	140, 1257	<0.0001	52.63
Species	28.19	8, 630	<0.0001	18.46
Species x Range	9.73	8, 630	<0.0001	7.48
Range	5.77	4, 315	0.0002	6.83
B. Limia				
Population	4.96	8, 62	<0.0001	28.11
Range	5.58	4, 31	0.0017	41.87
C. Poecilia				
Population	16.69	40, 574	<0.0001	62.28
Species	43.94	12, 400	<0.0001	52.37
Species x Range	8.23	12, 400	<0.0001	17.7
Range	3.62	4, 151	0.0075	8.76
D. Xiphophorus				
Population	3.35	24, 354	<0.0001	21.19
Species	19.89	8, 202	<0.0001	32.11
Species x Range	10.82	8, 202	<0.0001	21.17
Range	5.15	4, 101	0.0008	16.94
E. All Genera Combined				
Population	8.58	212, 2438	<0.0001	42.67
Species	35.18	28, 2201	<0.0001	28.36
Genus	70.21	12, 1614	<0.0001	30.95
Species x Range	9.45	28, 2201	<0.0001	12.75
Genus x Range	5.8	12, 1614	<0.0001	3.65
Range	7.52	4, 610	<0.0001	4.70

Note: MANCOVA results for life history analysis. F statistics p –values were calculated using JMP. Partial eta squares were calculated using the Wilkes Lambda calculation from Riesch *et al.* 2013.

Table 3.2 Analysis of Covariance for Life History Traits

Effect	F	DF	p
A. Fecundity			
Lot	13.8993	53, 613	<0.0001
Species	38.108	7, 613	<0.0001
Genus	26.0114	3, 613	<0.0001
Range	22.0793	1, 613	<0.0001
Species x Range	17.7416	7, 613	<0.0001
Genus x Range	11.9174	3, 613	<0.0001
B. Reproductive Weight			
Lot	13.1499	53, 613	<0.0001
Species	43.8502	7, 613	<0.0001
Genus	46.1369	3, 613	<0.0001
Range	14.0317	1, 613	0.0002
Species x Range	17.9925	7, 613	<0.0001
Genus x Range	10.1163	3, 613	<0.0001
C. Reproductive Allotment			
Lot	8.5816	53, 613	<0.0001
Species	17.084	7, 613	<0.0001
Genus	51.6501	3, 613	<0.0001
Range	9.1379	1, 613	0.0026
Species x Range	10.7787	7, 613	<0.0001
Genus x Range	5.1879	3, 613	0.0015
D. Average Embryo Weight			
Lot	4.3216	53, 613	<0.0001
Species	34.5647	7, 613	<0.0001
Genus	14.7371	3, 613	<0.0001
Range	0.7821	1, 613	0.3769
Species x Range	2.7394	7, 613	0.0083
Genus x Range	3.6491	3, 613	0.0125
E. Body Size			
Lot	12.7641	53, 613	<0.0001
Species	87.3349	7, 613	<0.0001
Genus	226.3429	3, 613	<0.0001
Range	6.1749	1, 613	0.0132
Species x Range	18.1664	7, 613	<0.0001
Genus x Range	10.626	3, 613	<0.0001

Note: Shows individual ANCOVA analyses for each life history trait of interest. A. Fecundity B. Reproductive Weight C. Reproductive Allotment D. Average Embryo Weight E. Body Size.

Appendix B

Table S.3.1 Collection Numbers and Locations

Genus	Species	Range	Populations Dissected
Gambusia	affinis	Invasive	9
Gambusia	holbrooki	Invasive	8
Gambusia	geiseri	Invasive	3
Limia	vittata	Invasive	2
Poecilia	latipinna	Invasive	2
Poecilia	mexicana	Invasive	2
Poecilia	reticulata	Invasive	3
Poecilia	vivipara	Invasive	3
Xiphophorus	helleri	Invasive	2
Xiphophorus	maculatus	Invasive	2
Xiphophorus	variatus	Invasive	2
Gambusia	affinis	Native	9
Gambusia	holbrooki	Native	10
Gambusia	geiseri	Native	1
Limia	vittata	Native	2
Poecilia	latipinna	Native	2
Poecilia	mexicana	Native	2
Poecilia	reticulata	Native	2
Poecilia	vivipara	Native	2
Xiphophorus	helleri	Native	2
Xiphophorus	maculatus	Native	2
Xiphophorus	variatus	Native	2

Note: This table shows the genera and species dissected for this project as well as the range they were sampled from and how many populations were dissected for each.

Table S.3.2 Descriptive statistics of Invasive and Non-invasive Poeciliid Life History

Invasion Status, Species	N	SL (mm)	TWW (g)	F	RWW (mg)	RA (%)	EW (mg)
Invasive Range							
G. affinis	82	37.80±5.1	1.12±.53	63.8±45.77	231.76±21.02	18.51±9.3	3.54±2.04
G. geiseri	30	34.09±3.35	.68±.2	16.83±8.06	100.78±7.6	15.32±4.73	6.25±2.16
G. holbrooki	68	34.79±4.91	.8±.29	36.55±19.03	175.31±10.3	21.7±8.51	4.8±1.97
Gambusia	180	36.04±5.02	.92±.45	45.63±37.55	188.53±10.91	19.2±8.66	4.47±2.25
L. vittata	21	57.29±7.67	4.6±1.82	87.71±37.64	616.09±81.36	13.19±5.35	6.7±3.08
Limia	21	57.29±7.67	4.6±1.82	87.71±37.64	616.09±81.36	13.19±5.35	6.7±3.08
P. latipinna	18	43.56±11.31	2.61±2.2	40.56±37.22	416.99±93.11	15.22±6.55	9.75±2.74
P. mexicana	16	40.39±8.68	1.90±.87	32.06±24.92	128.16±21.15	6.72±3.04	4.73±3.53
P. reticulata	32	30.73±4.56	.65±.32	27.66±19.7	94.25±12.42	13.43±5.43	3.67±2.58
P. vivipara	28	34.28±7.85	1.05±.68	33.29±19.81	106.86±13.57	10.92±4.72	3.16±1.4
Poecilia	94	35.89±9.23	1.35±1.32	32.55±24.85	165.58±22.62	11.88±5.78	4.86±3.49
X. helleri	17	45.76±5.09	1.81±.61	26.88±13.4	121.68±19.05	6.51±2.79	3.99±1.2
X. maculatus	18	38.41±5.17	1.68±.90	39.28±18.89	182.68±31.19	11.03±6.17	3.83±1.74
X. variatus	19	38.61±4.45	1.43±.48	51.26±18.17	229.59±21.51	16.2±4.51	4.4±1.54
Xiphophorus	54	40.8±5.89	1.63±.69	39.59±19.53	179.98±15.18	11.43±6.12	4.08±1.51
Total/Avg	350	38.01±8.44	1.37±1.25	43.71±34.43	206.7±11.29	15.67±8.27	4.65±2.66
Native Range							
G. affinis	83	34.77±3.94	.78±.29	44.17±21.43	145.93±10.58	17.68±8.5	2.89±1.38
G. geiseri	10	35.64±3.63	.79±.25	27.6±9.79	125.76±21.56	17.24±5.24	4.76±1.82
G. holbrooki	86	31.49±5.38	.58±.32	19.72±14.02	85.54±7.5	14.55±6.83	4.02±1.79
Gambusia	179	33.24±4.95	.68±.32	31.5±21.29	116.51±6.57	16.15±7.7	3.54±1.73
L. vittata	18	53.83±7.81	3.84±1.62	46.28±18.84	293.02±47.03	7.71±4.09	6.1±3.89
Limia	18	53.83±7.81	3.84±1.62	46.28±18.84	293.02±47.03	7.71±4.09	6.1±3.89
P. latipinna	20	37.35±2.78	1.27±.3	26.25±12.86	274.25±31.26	20.67±7.75	10.14±3.48
P. mexicana	16	56.33±11.24	4.35±2.19	54.13±41.66	350.73±53.53	8.31±3.09	7.08±4.24
P. reticulata	26	30.17±3.9	.63±.24	23.88±11.4	54.15±5.43	8.95±3.83	2.2±.73
P. vivipara	17	32.55±6.51	.86±.49	24.65±17.97	105.31±19.54	11.00±5.04	4.21±2.48
Poecilia	79	37.8±11.64	1.59±1.74	30.77±24.96	180.94±29.44	12.23±7.19	5.63±4.23
X. helleri	21	53.09±6.2	2.94±.95	47.48±21.58	370.66±44.52	12.07±5.41	7.33±3.64
X. maculatus	16	27.28±2.87	.55±.27	9.5±4.62	57.08±10.41	9.89±4.62	4.32±2.59
X. variatus	26	36.03±5.54	1.08±.46	31.35±20.01	172.23±17.3	15.64±4.28	5.99±2.33
Xiphophorus	63	39.49±11.53	1.57±1.18	31.17±22.98	209.13±22.66	12.99±5.26	6.01±3.07
Total/Avg	339	36.56±9.76	1.23±1.31	32.05±22.55	157.99±7.93	14.2±7.38	4.62±3.08

Note: Means and standard deviations are shown for female life history traits. Number of individuals dissected (N), Standard Length of the adult in mm (SL), Total wet weight of the adult in g (TWW), Fecundity (F), Reproductive allotment, Reproductive weight/total wet weight given as a % (RA), Average embryo weight in mg (EW).

Table S.3.3 Partial Correlations of Life History Traits

A.

	Fecundity	Embryo Weight	Reproductive Allotment	Body Size
Fecundity	1	0.1759	<0.0001	<0.0001
Embryo Weight	-0.9085	1	<0.0001	<.0001
Reproductive Allotment	0.9238	0.9258	1	0.3368
Body Size	0.9502	0.917	-0.9078	1

B.

	Fecundity	Embryo Weight	Reproductive Weight	Body Size
Fecundity	1	0.1759	<0.0001	<0.0001
Embryo Weight	-0.9085	1	<0.0001	<0.0001
Reproductive Weight	0.9238	0.9258	1	<0.0001
Body Size	-0.0429	0.1223	0.3192	1

Note: values are R^2 correlation coefficients, below the diagonal, generated from Multivariate correlation analysis in JMP. P values are above the diagonal. Table A shows correlations with Reproductive Allotment; Table B shows correlations with Reproductive Weight.

Table S.3.4 Collection Information

Museum	Lot #	Genus	Species	Country	Location	Year
UMMZ	81123	Gambusia	affinis	USA	Oklahoma	1927
UMMZ	92379	Gambusia	affinis	USA	Texas	1931
UMMZ	115159	Gambusia	affinis	USA	Alabama	1936
UMMZ	121655	Gambusia	affinis	USA	New Mexico	1937
UMMZ	123134	Gambusia	affinis	USA	Arkansas	1936
UMMZ	125010	Gambusia	affinis	USA	Nevada	1936
UMMZ	140419	Gambusia	affinis	USA	California	1942
UMMZ	141396	Gambusia	affinis	USA	Utah	1942
UMMZ	141847	Gambusia	affinis	USA	Arizona	1944
UMMZ	162644	Gambusia	affinis	USA	California	1959
UMMZ	163631	Gambusia	affinis	USA	Mississippi	1941
UMMZ	170527	Gambusia	affinis	USA	Louisiana	1940
UMMZ	184184	Gambusia	affinis	USA	Louisiana	1956
UMMZ	197975	Gambusia	affinis	Guam	Mariana Islands	1961
UMMZ	203020	Gambusia	affinis	Mexico	Chihuahua	1978
UMMZ	204203	Gambusia	affinis	Bolivia	Cochabamba	1964
UMMZ	210406	Gambusia	affinis	USA	Texas	1940
UMMZ	220392	Gambusia	affinis	USA	Nebraska	1989
USNM	246459	Gambusia	affinis	Fiji	NA	NA
USNM	166864	Gambusia	affinis	Egypt	NA	NA
TNHC	39533	Gambusia	clarkhubbsi	USA	Texas	2003

TNHC	45363	Gambusia	clarkhubbsi	USA	Texas	2008
TNHC	39522	Gambusia	clarkhubbsi	USA	Texas	2007
TNHC	39526	Gambusia	clarkhubbsi	USA	Texas	2004
TNHC	7030	Gambusia	geiseri	USA	Texas	1961
TNHC	21996	Gambusia	geiseri	USA	Texas	1990
TU	24783	Gambusia	geiseri	USA	Texas	1961
UMMZ	120358	Gambusia	geiseri	USA	Texas	1938
UMMZ	132270	Gambusia	geiseri	USA	Texas	1940
ANSP	126267	Gambusia	holbrooki	Bermuda	Hamilton Island	1973
ANSP	168877	Gambusia	holbrooki	Bermuda	NA	1985
UMMZ	166646	Gambusia	holbrooki	Egypt	NA	1953
UMMZ	56193	Gambusia	holbrooki	USA	Georgia	1922
UMMZ	66941	Gambusia	holbrooki	USA	Florida	1925
UMMZ	88397	Gambusia	holbrooki	USA	Georgia	1929
UMMZ	88558	Gambusia	holbrooki	USA	Florida	1929
UMMZ	88784	Gambusia	holbrooki	USA	Alabama	1929
UMMZ	114101	Gambusia	holbrooki	USSR	Uzbekistan	1936
UMMZ	121616	Gambusia	holbrooki	Italy	Padova	1939
UMMZ	138464	Gambusia	holbrooki	USA	Maryland	1939
UMMZ	14413	Gambusia	holbrooki	USSR	Uzbekistan	1936
UMMZ	88503	Gambusia	holbrooki	USA	Georgia	1931
UMMZ	126283	Gambusia	holbrooki	USA	North Carolina	1930
UMMZ	166655	Gambusia	holbrooki	Egypt	NA	1953
UMMZ	201611	Gambusia	holbrooki	India	Kashmir	1977
USNM	88476	Gambusia	holbrooki	USA	NA	NA
UMMZ	124137	Gambusia	holbrooki*	South Africa	NA	1939
UMMZ	200770	Gambusia	holbrooki*	USA	Florida	1977
UMMZ	210129	Gambusia	holbrooki*	USA	Florida	1966
UMMZ	179801	Gambusia	nobilis	USA	Texas	1961
TNHC	25186	Gambusia	speciosa	USA	Texas	1997
TNHC	27490	Gambusia	speciosa	USA	Texas	1999
TNHC	27496	Gambusia	speciosa	USA	Texas	1999
TNHC	29074	Gambusia	speciosa	USA	Texas	1997
UMMZ	130352	Gambusia	speciosa	Mexico	Coahuila	1939
UMMZ	97549	Gambusia	speciosa	Mexico	Nuevo Leon	1930
UMMZ	120320	Gambusia	speciosa	USA	Texas	1936
UMMZ	196738	Gambusia	speciosa	Mexico	Coahuila	1974
TNHC	43520	Gambusia	vittata	Mexico	Tamaulipas	1983
TU	43604	Gambusia	vittata	Mexico	San Luis	1960
UMMZ	97511	Gambusia	vittata	Mexico	Tamaulipas	1939
TNHC	45666	Gambusia	clarkhubbsi	USA	Texas	2006
UF	118806	Limia	tridens	Haiti	Hispaniola	1951
UF	118807	Limia	tridens	Haiti	Hispaniola	1949

UF	92406	Limia	vittata	Cuba	Guantanamo	1943
UF	92418	Limia	vittata	Cuba	Santiago de Cuba	1948
UMMZ	194206	Poecilia	gilli	Costa Rica	San Jose	1973
UMMZ	173233	Poecilia	gilli	Honduras	Santa Barbara	1951
UMMZ	197483	Poecilia	gilli	Fiji	NA	1970
USNM	293483	Poecilia	gilli	Panama	NA	NA
ANSP	84089	Poecilia	latipinna	USA	Georgia	1941
ANSP	95088	Poecilia	latipinna	USA	Texas	1960
UMMZ	203274	Poecilia	latipinna	Mexico	Tamaulipas	1965
UMMZ	203281	Poecilia	latipinna	Mexico	Tamaulipas	1965
UMMZ	95812	Poecilia	latipinna	USA	Mississippi	1933
UMMZ	196866	Poecilia	latipinna	USA	Hawaii	1974
UMMZ	213906	Poecilia	latipinna	USA	Hawaii	1951
TU	96134	Poecilia	latipunctata	Mexico	Tamaulipas	1964
UMMZ	97695	Poecilia	latipunctata	Mexico	Tamaulipas	1930
UMMZ	189555	Poecilia	mexiana	USA	Nevada	1964
TNHC	27261	Poecilia	mexicana	Mexico	Tamaulipas	1999
TU	43593	Poecilia	mexicana	Mexico	Tamaulipas	1960
TU	43602	Poecilia	mexicana	Mexico	San Luis	1960
TU	84696	Poecilia	mexicana	Costa Rica	Puntarenas	1962
UF	27602	Poecilia	mexicana	Panama	San Blas	1965
UMMZ	189544	Poecilia	mexicana	USA	Nevada	1963
UMMZ	143639	Poecilia	petenensis	Guatemala	Peten	1935
UMMZ	192815	Poecilia	petenensis	Mexico	Tobasco	1968
UMMZ	143570	Poecilia	petenensis	Guatemala	NA	1935
UMMZ	143632	Poecilia	petenensis	Guatemala	Peten	1935
UMMZ	184722	Poecilia	petenensis	Mexico	Tobasco	1959
UMMZ	196613	Poecilia	petenensis	Mexico	Campeche	1974
UF	74903	Poecilia	picta	Trinidad	NA	1973
TU	30815	Poecilia	reticulata	Mexicos	Michoacan	1963
TU	123712	Poecilia	reticulata	USA	Nevada	1981
UF	30397	Poecilia	reticulata	Dominican Republic	Hispaniola	1977
UF	91918	Poecilia	reticulata	USA	Whoming	1984
UF	178173	Poecilia	reticulata	Thailand	Chiang Mai	2010
UMMZ	172641	Poecilia	reticulata	Puerto Rico	NA	1954
UMMZ	215283	Poecilia	reticulata	Brazil	Rio de Janeiro	1988
UMMZ	143098	Poecilia	velifera	Mexico	Yucatan	1936
UMMZ	143099	Poecilia	velifera	Mexico	Yucatan	1936
UMMZ	196582	Poecilia	velifera	Mexico	Campeche	1974
ANSP	174188	Poecilia	vivipara	Brazil	Bahia	1995
ANSP	174200	Poecilia	vivipara	Brazil	Espirito Santo	1995
UMMZ	172750	Poecilia	vivipara	Puerto Rico	NA	1955
UMMZ	198774	Poecilia	vivipara	Puerto Rico	NA	1974

UMMZ	198776	Poecilia	vivipara	Puerto Rico	NA	1973
UMMZ	203416	Poecilia	vivipara	Brazil	Ceara	1935
USNM	101446	Poecilia	vivipara	Puerto Rico	NA	NA
UMMZ	214879	Limia	vittata	USA	Hawaii	1951
USNM	247416	Limia	vittata	USA	NA	NA
UMMZ	214879	Limia	vittata	USA	Hawaii	1951
ANSP	75876	Poecilida	picta	Trinidad	La Brea	1930
UF	118152	Xiphophors	maculatus	USA	Florida	1981
TU	94348	Xiphophorus	helleri	USA	Nevada	1975
UMMZ	193914	Xiphophorus	helleri	Guatemala	Alta Verapaz	1973
UF	91919	Xiphophorus	hellerii	USA	Nevada	1980
UF	171132	Xiphophorus	hellerii	USA	Florida	1972
UF	171349	Xiphophorus	hellerii	USA	Florida	1970
UMMZ	97578	Xiphophorus	hellerii	Mexico	Veracruz	1930
UMMZ	172640	Xiphophorus	maculatus	Puerto Rico	NA	1954
UMMZ	143780	Xiphophorus	maculatus	Guatemala	Peten	1935
UMMZ	202847	Xiphophorus	maculatus	Belize	Orange Walk	1954
USNM	245937	Xiphophorus	maculatus	USA	NA	NA
UMMZ	177310	Xiphophorus	milleri	Mexico	Veracruz	1957
UMMZ	184556	Xiphophorus	milleri	Mexico	Veracruz	1959
TNHC	41433	Xiphophorus	variatus	USA	Texas	2008
UF	99006	Xiphophorus	variatus	USA	Florida	1993
UMMZ	97575	Xiphophorus	variatus	Mexico	Veracruz	1930
UMMZ	194929	Xiphophorus	variatus	Mexico	Veracruz	1948
UMMZ	108662	Xiphophorus	xiphidium	Mexico	Tamaulipas	1930
UMMZ	108663	Xiphophorus	xiphidium	Mexico	Tamaulipas	1930
UMMZ	124416	Xiphophorus	xiphidium	Mexico	Tamaulipas	1939
UMMZ	162151	Xiphophorus	xiphidium	Mexico	Tamaulipas	1940

Note: Table S.3.4 shows the collection information for the individuals used for chapters 2 and 3 of this dissertation. Musuem, Lot Number, Genus, Species, Country of origin, Location, and Year of collection are given. *Holbrooki** were originally identified as *Gambusia affinis* but upon inspection were found to be *Gambusia holbrooki.*

CHAPTER 4: SEQUENCING OF A HIGH QUALITY REFERENCE GENOME FOR THE INVASIVE MOSQUITOFISH *GAMBUSIA AFFINS* USING ILLUMINA SHORT READ SEQUENCING IN CONJUNCTION WITH A CHICAGO LIBRARY[3]

[3] Troendle NJ Glenn T Mahmud O Chalopin D and Mauricio R. To be submitted to *Plos One*

Abstract

Invasive species are an issue of great concern around the world and are responsible for extensive ecological and economic costs. The invasive western mosquitofish, *Gambusia affinis,* is one of the most invasive fish species in the world and, together with its sister species *Gambusia holbrooki*, has successfully invaded over 50 countries on six continents. In order to better understand how this species became so invasive, we need tools to perform genomic studies. This paper outlines the creation of a high quality reference genome for *Gambusia affins*. We used Illumina short read sequencing of traditional paired end libraries in conjunction with the new Chicago Libraries. We assembled a reference genome with high contiguity and coverage, with N50 contig and N50 scaffold lengths of 17.6Kb and 6.65Mb respectively and total estimated coverage of 54.9X. We then annotated the genome and compared its quality to three other fish genomes to ensure that our assembly was of high quality.

Introduction

The invasive mosquitofish *Gambusia affinis* is considered to be one of the most widespread freshwater fish in the world. In the last century it has rapidly expanded from its native range in the southeastern United States and Northern Mexico and has become established in more than 50 countries and every continent besides Antarctica (Krumholz, 1948; Lever, 1996). Originally introduced to control mosquito populations during malarial and yellow fever outbreaks this small fish has had severe impacts on the native fauna in its introduced range leading to declines in insect, fish and amphibian populations (Pyke, 2008). We are interested in understanding what makes this species and others like it such successful invaders while other species struggle even within their native ranges. To do this we need a detailed understanding not only of the phenotypic traits that these species exhibit but the underlying genetic behind those traits.

With the falling cost of genomic sequencing over the past decade it has become more feasible for individual labs to sequence genomes. The sequencing of a genome opens up many possibilities for genomic studies. Many non-model organisms are limited in the types of studies that can be done because they have no sequenced genome. The creation of a high quality genome assembly of the invasive species *Gambusia affinis* would be an excellent resource for studying invasive species as well as those studying other aspects of the family Poeciliidae. At the time of this project there are only two members of the family Poeciliidae with sequenced genomes, the platyfish, *Xiphophorus maculatus* (Schartl *et al.* 2013), and the guppy, *Poecilia reticulata* (Fraser *et al.* 2011). The creation of a third genome from a third genus will allow greater comparisons of evolutionary history within the family. It would allow researchers to make genomic comparisons between these three species. In addition this assembly could be used

as a reference genome for assembling other *Gambusia* genomes and conducting comparative genomic studies.

The family Poecilidae has been a group of interest and has been highly studied. Researchers have used this family to contribute to the fields of life history evolution, phenotypic plasticity, sex-chromosome evolution and invasion (Reznick and Endler, 1982; Trexler and Travis, 1990; Sakai *et al.* 2001, Lamatsch, *et al.* 2015). Reznick's work on predator prey interactions in the guppy, *Poecilia reticulata* has become one of the primary cases in life history theory. Poeciliidae are also unique in that the family has several different sex-determination methods, with multiple methods existing even within a single genus or species (Volf and Schartl, 2001). The family Poeciliidae is viviparous and has been used as a model for studying the evolution of the placenta and viviparity (Pollux *et al.* 2009). Poeciliidae have also been used as models of sexual selection, genital evolution, adaptations to extreme environments, behavior, and genetics (Evans *et al.* 2011).

Beyond creating a general resource, our goal was to create a tool that would allow us to investigate the evolutionary history and genetic basis of specific traits that are associated with invasive success. We also wished to investigate the degree of genetic diversity in the invasive range of *G. affinis*. Invasive species are often expected to have decreased genetic diversity due to genetic bottlenecks and founder's events (Nei *et al.* 1975). However, recent studies have found that this is not always the case and in some situations multiple introductions may lead to increased diversity in the invasive range (Dlugosch and Parker, 2008).

To make these types of analyses possible, we have created a high quality genome assembly using Illumina sequencing of both traditional paired end libraries and a new mate pair alternative Chicago library (Putnam et al, 2016) which is designed to increase contiguity of

genome assemblies sequenced using short read data. This method has been shown to increase assembly contiguity by up to twenty-fold while still using short read data from the Illumina HiSeq platform. One measure of a genomes quality is the contiguity. The more contiguous a genome is the fewer breaks there are in sequencing which means that genes are less likely to be split up and we have a better understanding of how the whole genome fits together. While short read sequencing is much less expensive than traditional Sanger sequencing, the read lengths are much shorter resulting in a more fragmented genome assembly (Liu *et al.* 2012) which cannot be overcome simply by increasing sequencing depth due to the repetitive sequences found within genomes. Traditional mate-pair libraries are intended to bridge the gaps caused by repetitive sequences however they are often not long enough to do so for genomes with a large proportion of repetitive DNA (Putnam *et al.* 2016). The Chicago library method in conjunction with the HiRise software pipeline is designed to bridge these gaps of repetitive sequences (Putnam *et al.* 2016). Using this method, we have assembled a high quality genome with a scaffold N50 of 914KB. We have also annotated this genome using genes pulled from the *Xiphophorus maculatus* and *Poecilia reticulata* genomes. Finally we made comparisons between the genomes of *Poecilia reticulata, Oryzias latipes, Xiphophorus maculatus,* and *Gambusia affinis* to compare the degree of synteny and proportion of the genome composed of transposable elements, in order to determine if our genome assembly falls within the normal range for a Poeciliid genome.

Methods

Source Material

Two fish were used to create our genome assembly. For the short insert library, we extracted DNA from the muscle tissue of a male *Gambusia affinis* from the Zuibaiji river in Japan using a modified phenol-chloroform extraction. For the Chicago library, we extracted DNA from muscle tissue of a male *G. affinis* from the Platt river in Nebraska. Male fish were chosen for this because males possess homomorphic sex chromosomes in the species *G. affinis* (Volf and Schartl, 2001).

Illumina library preparation and sequencing

We sheared the genomic DNA using the Covaris S2 targeting a 600bp average fragment size. The sheared DNA was end-repaired, adenylated and ligated to TruSeq LT adapters using an Illumina PCR free library preparation kit. We purified the ligation reaction using a Qiagen Qiaquick Gel extraction Kit from a 2% agarose gel. We sequenced the library on a HiSeq 2500 to obtain paired-end 100 base reads.

De Novo shotgun assembly using Meraculous

We produced a draft genome assembly for *Gambusia affinis* using a modified version of Meraculous 2.0 (Chapman et al, 2011) with a kmer size of 31bp for *de novo* assembly of 32x coverage in paired-end reads using a full run of Illumina HiSeq. We trimmed reads for quality and sequencing adapters using Trimmomatic (Bolger et al 2015). We then used the resulting

594.6Mbp assembly, with scaffold N50 of 31kbp and contig N50 of 13.9kbp, as the input assembly for HiRise scaffolding.

Chicago library prep

We prepared a Chicago library as described previously (Putnam et al, 2016). We extracted ≥ 0.5 μg of high molecular weight genomic DNA (~50 kbp mean fragment size) from a single individual, reconstituted into chromatin *in vitro*, and fixed with formaldehyde. We then digested fixed chromatin with DpnII, filled in the 5' overhangs with biotinylated nucleotides, and ligated the free blunt ends. After ligation, we reversed crosslinks and purified the DNA from protein. We then treated the purified DNA to remove biotin that was not internal to ligated fragments. We sheared the DNA to ~350 bp mean fragment size, and generated sequencing libraries using NEBNext Ultra enzymes and Illumina-compatible adapters. We then isolated biotin-containing fragments using streptavidin beads before PCR enrichment of the library. The Chicago library was also sequenced on a HiSeq 2500 to obtain paired-end 100 base reads.

Scaffolding the draft genome with HiRise

We used the *Gambusia affinis* draft genome in FASTA format (described above), shotgun sequences, and Chicago library sequence (57M read pairs; 2X125bp) in FASTQ format as input data for HiRise, a software pipeline designed specifically for using Chicago library sequence data to assemble genomes (Putnam et al, 2016). We aligned the Shotgun and Chicago library sequences to the draft input assembly using a modified SNAP read mapper (http://snap.cs.berkeley.edu). We analyzed the separations of Chicago read pairs mapped within draft scaffolds by HiRise to produce a likelihood model, and used the resulting likelihood model

to identify putative misjoins and score prospective joins. After this we used scaffolding, shotgun sequences to close gaps between contigs.

Gene prediction and annotation of the Gambusia affinis genome

We used the MAKER genome annotation pipeline to generate *G. affinis* gene annotations (Campbell *et al.* 2014). MAKER combines several classes of evidence data – for example RNAseq data or proteins from closely related species - to generate ab initio gene predictions by using several tools within its pipeline. Our MAKER pipeline consisted of the following steps: A) RNAseq and protein sequences from *Xiphophorus maculatus* (Schartl *et al.* 2013) and *Poecilia reticulata* (Fraser *et al.* 2011) were used as evidence for the initial annotations; B) The initial annotations were used to train SNAP gene prediction tool multiple times (Korf, 2004); C) The final set of gene annotations were generated from the trained *ab initio* SNAP predictions.

To assess the quality of *G. affinis* gene annotations, we used BLAST (e-value: 1e-10) to compare them to those of *Xiphophorus maculatus*, *Poecilia reticulata*, and *Oryzias latipes* (Kasahara *et al.* 2007). These organisms were chosen because *X. maculatus* and *P. reticulata* are the only two other Poeciliidae that have fully sequenced genomes and *O. latipes* has a high quality genome and is often used for comparisons in fish species.

To functionally annotate *G. affinis* genes, putative functions were assigned to gene annotations using BLASTP (e-value: 1e-20) to identify the best homologs from the UniProt/Swiss-Prot protein database (Pundir *et al.* 2016). InterProScan was used to find and assign protein domains to gene annotations (Quevillon *et al.* 2005).

Orthology Clustering

To find *G. affinis* species-specific genes and expanded gene families, we grouped protein sequences of *G. affinis*, *X. maculatus,* and *O. latipes* into orthologs and paralogs using OrthoMCL (Li *et al.* 2003) [OrthoMCL Parameters: Inflation: 1.5; P-value cutoff: 1e-05; P-ident cutoff: 0; P-match cutoff: 0; Maximum weight: 350].

Transposable Elements

To determine the repeat content in the assembly, we used RepeatModeler (Smit, AFA, Hubley, R. *RepeatModeler Open-1.0*) with default parameters to identify and build an automatic library containing interspersed elements (transposable elements), simple repeats and low complexity regions. This library contains 737 consensus sequences not detected with RepeatModeler, we classified Miniature Inverted Repeat Elements (MITEs) using MITE-Hunter (Han and Wessler 2010). The program found 170 consensus elements, from which we classified 102 in 24 families and 68 as singlet families. To filter false positive, we verified various MITE-specific criteria: TIR and TSD can be determined on multiple sequence alignment (MSA) while flanking regions are divergent; MITEs are usually high copy number in their host genome; the autonomous associated DNA transposons can sometimes be identified. Thus, we performed the following analyses for each consensus sequence: observation of MSA file, use of BlastN2 and RNAfold (Lorenz *et al.* 2011) to help identifying the TIRs and TSDs, family identification using CENSOR (Kohany *et al.* 2006), and copy number estimation in the assembly via Blast analysis. At last, 35 sequences were added in the repeat library. 20 of these sequences were found to be very conserved in *Xiphophorus maculatus* genome (full length sequence, >90% of identity).

We masked he assembly using RepeatMasker 4.0.0 (Smit, AFA, Hubley, R & Green, P. *RepeatMasker Open-4.0*) with the *G.affinis*-specific library and the -lib option. For each classes and transposable element families, we estimated the number of copies as well as the coverage in the assembly from the RepeatMasker outfile ".out".

Synteny

We performed pairwise analyses of synteny between *G. affinis* and three other fish species. We used the two other sequenced Poeciliidae, *Xiphophorus maculatus* and *Poecilia reticulata* and the *Oryzias latipes* (medaka) genome to investigate co-linearity of our genome when compared to these other species (Figure 4.1). We used CoGe Synmap (Lyons *et al.* 2008), to perform the analyses mapping the genome with the greater number of scaffolds to the one with fewer. Only scaffolds larger than 1GB were used in these analyses.

Results

Assembly

We sequenced the whole genome of the mosquitofish *Gambusia affinis* using one male fish from the Zuibaji River region in Japan for the initial shotgun sequencing and a second male fish from the Platt river system in Nebraska, USA for the HiRise sequencing. We produced a genome assembly with 54.9x coverage and an N50 contig and scaffold size of 17.6KB and 6.65MB respectively (Table 4.1). This was a large improvement over the initial shotgun assembly, which had contig and scaffold sizes of 13.9KB and 31KB (Figure 4.1). In addition to the increase in scaffold size we also had a large increase in contiguity with the number of

scaffolds in the N50 decreasing from 5240 in the Meraculous assembly to 26 in the HiRise assembly.

Annotation

The final annotation set of the *Gambusia affinis* genome from the MAKER annotation pipeline contained 21163 predicted genes. BLASTp analyses revealed 20511 (97%), 19904 (94%) and 18880 (89%) of the predicted genes had significant hits to *P. reticulata, X. maculatus* and *O. latipes* respectively. Average gene, exon and intron lengths are relatively smaller in *G. affinis* when compared to closely related organisms (*P. reticulata, X. maculatus* and *O. latipes*), but average coding sequence length and exons per gene are comparable (Table 4.1).

Transposable Elements

The assembly contains about 20% of repeats with 17.7% of transposable elements (Table 4.2). Among transposable elements, the DNA transposon is the most abundant class and in particular TcMariner and hAT families. The *G. affinis* assembly seems to be less repetitive than other sequenced poeciliid genomes (from the *Xiphophorus* genus, Shen *et al.* 2016), which harbor a higher content of TcMariner and hAT families.

Synteny

We found that there were large sections of our genome that were highly syntenic with the genomes of the two other Poeciliid species *X. maculatus* and *P. reticulata* and had very few rearrangements (Figure 4.2). We also compared each of the Poeciliid genomes to the genome of *Oryzias latipes* (Medaka), which is one of the most complete fish genomes. We found that there

were higher degrees of synteny between the three Poeciliidae species than when they were compared to Medaka. We also observed that greater synteny between *G. affinis* and *X. maculatus* than between *G. affinis* and *P. reticulata* while synteny between *X. maculatus* and *P. reticulata* was intermediate between the two.

Discussion

We sequenced and assembled the genome of the Mosquitofish *Gambusia affinis* using short read illumina sequencing of paired end and Chicago libraries. The resulting genome assembly has excellent coverage, scaffold sizes and contiguity. A total of 598.7MB of the estimated 733.5-919.3MB was sequenced (Lamatsch *et al.* 2000). Using the new Chicago library our assembly's N50 scaffold size was larger and composed of fewer scaffolds than other assemblies that utilized Roche 454 long insert sequencing like the *Xiphophorus* genome (Shartl *et al.* 2013). The result is a high quality genome composed of large pieces with some scaffolds approaching estimated chromosome size.

Our genome annotation revealed that *G. affinis* has a similar number of genes when compared to its close relatives, *X. maculatus*, *P. reticulata.* We also found that the number of genes did not significantly differ from *Oryzias latipes,* which is much more distantly related. In addition, 94%, 97% and 89% of the genes found in *G. affinis* had hits in the *X. maculatus, P. reticulata,* and *O. latipes* genomes respectively indicating that many of these genes are conserved along this lineage. These results suggest that our genome assembly contains the majority of the genes that have been identified in the both in the Poeciliid family and in other fish species.

We also found that the genome of *G. affinis* is composed of about 20% repetitive elements, 17.7% of the genome is made up of transposable elements. This number is only slightly lower than other Poeciliidae genomes (Shen *et al.* 2016).

We conducted pairwise comparisons of synteny between *G. affinis, X. maculatus, P. reticulata,* and *O. latipes* and found that within the family Poeciliidae the genomes are highly syntenic with only minor inversions and translocations. We did find that *X. maculatus*, which is between *G. affinis* and *P. reticulata* phylogenetically, showed greater synteny with both of these species than either did with each other. When comparing each of the Poeciliid species to Medaka, which is distantly related, we still found high degrees of synteny. However, *P. reticulata,* which is the most closely related of the Poeciliidae, had far more rearrangements than the other two species did. This could be due to the larger scaffold size of the *P. reticulata* genome. Because SynMap aligns the two genomes together assemblies composed of smaller pieces may align better than ones composed of very large pieces which would make it appear that there is greater synteny in the genomes made from smaller pieces.

The genome assembly that we have created will be a useful resource not only for our lab and the study of invasive species but for other Poeciliid researchers as well. We hope to use it to shed light on the evolution of invasiveness in Poeciliid fishes. Having a high quality genome assembly will allow researchers to perform analyses that were not possible before now. It is our hope that the creation of this genome will provide greater understanding of Poeciliid evolution.

References

Bolger A.M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114-20.

Campbell, M.S., C. Holt, B. Moore, and M. Yandell. 2014. Genome annotation and curation using maker and Maker-P. Current Protocols in Bioinformatics **48**: 1-39.

Chapman J.A., I. Ho, S. Sunkara, S. Luo, G.P. Schroth, and D.S. Rokhsar. 2011. Meraculous: *De novo* genome assembly with short paired-end reads. PLoS One **6**: e23501.

Dlugosch K.M., and I.M. Parker. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology **17**: 431-49.

Evans, J.P., A. Pilastro, and I. Schlupp. 2011. Ecology and Evolution of Poeciliid Fishes. University of Chicago Press, Chicago, IL.

Fraser, B.A., C.J. Weadick, I. Janowitz, F.H. Rodd, and K.A. Hughes. 2011. Sequencing and characterization of the guppy (*Poecilia reticulata*) transcriptome. BMC Genomics **12**: 1-14.

Han, Y., and S.R. Wessler. 2010. MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acid Resources **38**: 1-8.

Kasahara, M., K. Naruse, S. Sasaki, Y. Nakatani, W. Qu, B. Ahsan, T. Yamada, Y. Nagayasu, K. Doi, Y. Kasai, T. Jindo, D. Kobayashi, A. Shimada, A. Toyoda, Y. Kuroki, A. Fujiyama, T. Sasaki, A. Shimizu, S. Asakawa, N. Shimizu, S. Hashimoto, J. Yang, Y. Lee, K. Matsushima, S. Sugano, M. Sakaizumi, T. Narita, K. Ohishi, S. Haga, F. Ohta, H. Nomoto, K. Nogata, T. Morishita, T. Endo, T. Shin-I, H. Takeda, S. Morishita, and Y. Kohara. 2007. The Medaka draft genome and insights into vertebrate genome evolution. Nature **447**: 714-719.

Kohany, O., A.J. Gentles, L. Hankus, and J. Jurka. 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics **7**: 1-7.

Korf, I. 2004. Gene finding in novel genomes. BMC Bioinformatics **5**: 1-7.

Lamatsch, D.K., S. Adolfsson, A.M. Senior, G. Christiansen, M. Pichler, Y. Ozaki, L. Smeds, M. Schartl, and S. Nakagawa. 2015. A Transcriptome derived female-specific marker from the invasive western mosquitofish (*Gambusia affinis*). PLoS One **10**: 1-16.

Lamatsch, D.K., C. Steinlein, M. Schmid, and M. Schartl. 2000. Noninvasive determination of genome size and ploidy level in fishes by flow cytometry: detection of triploid *Poecilia formosa*. Cytometry **39**: 91-95.

Li, L., C.J. Stoeckert Jr., and D.S. Roos. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Research **13**: 2178-2189.

Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, and M. Law. 2012. Comparison of next-generation sequencing systems. Journal of Biomedicine and Biotechnology **2012**: 1-11.

Lorenz, R., S.H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, and I.L. Hofacker. 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology **6**: 1-14.

Lyons E., B. Pedersen, J. Kane, and M. Freeling. 2008. The value of nonmodel genomes and an example using SynMap within CoGe to dissect the hexaploidy that predates Rosids. Tropical Plant Biology **1**: 181-190.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution **29**: 1-10.

Pollux, B.J.A., M.N. Pires, A.I. Banet, and D.N. Reznick. 2009. Evolution of placentas in the fish family Poeciliidae: An empirical study of macroevolution. Annual Review of Ecology, Evolution and Systematics **40**: 271-289.

Pundir, S., M.J. Martin, and C. O'Donovan. 2016. UniProt tools. Current Protocols in Bioinformatics **53**: 1-29.

Putnam N.H., B.L. O'Connell, J.C. Stites, B.J. Rice, M. Blanchette, R. Calef, C.J. Troll, A. Fields, P.D. Hartley, C.W. Sugnet, D. Haussler, D.S. Rokhsar, R.E. Green. 2016. Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. Genome Research **26**: 342-50.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, R. Apweller, and R. Lopez. 2005. InterProScan: Protein domains identifier. Nucleic Acid Research **33**: 116-120.

Shen, Y., D. Chalopin, T. Garcia, M. Boswell, W. Boswell, S.A. Shiryev, R. Agarwala, J.J. Volff, J.H. Postlethwait, M. Schartl, and P. Minx. 2016. *X. couchianus* and *X. hellerii* genome models provide genomic variation insight among *Xiphophorus* species. BMC Genomics **17**: 1-13.

Smit, A.F.A., R. Hubley, and P. Green. 2013-2015. RepeatMasker Open-4.0. Available at http://www.repeatmasker.org.

Trexler, J.C., and J. Travis. 1990. Phenotypic plasticity in the sailfin molly, *Poecilia latipinna* (Pisces: Poeciliidae). I. Field experiments. Evolution 143-156.

Volf, J.M., and M. Schartl. 2001. Variability of genetic sex determination in poeciliid fishes. Genetica **111**: 101-110.

Tables and Figures

Table 4.1. Genome assembly quality statistics

	Meraculous Assembly	Dovetail HiRise Assembly
Total Length	594.6 Mb	598.7 Mb
Scaffold N50	31 kb	6.65 Mb
Scaffold N90	7 kb	914 kb
Scaffold L50	5,240 scaffolds	26 scaffolds
Scaffold L90	20,613 scaffolds	117 scaffolds
Longest scaffold	324444	24339338
Number of scaffolds	38526	2943
Number of scaffolds >1 kb	38519	2940
Contig N50	13.9 kb	17.6 kb
Contig N90	3.56 kb	4.23 kb
Contig L50	12100 contigs	9490 contigs
Contig L90	44284 contigs	35674 contigs
Number of gaps >= 100 Ns	18145	40532
Percent of genome in gaps	0.972%	1.34%

Note: This figure shows the genome quality statistics for initial shotgun sequencing assembled by Meraculous and the final HiRise assembly.

Table 4.2: Summary of general statistics for the *Gambusia affinis* genome annotation

	G. affinis	*P. reticulata*[1]	*X. maculatus*[2]	*O. latipes*[3]
Number of protein encoding genes	21,144	22,982	22,082	22,658
Mean gene length (bp)	13,510	18,441	15,702	16,221
Mean CDS length (bp)	1,827	2,175	1,714	1,893
Exons				
Number	236,097	276,363	227,016	258,916
Mean length (bp)	164	267	189	260
Mean number per gene	11	12	10	11
Introns				
Number	214,953	248,065	205,251	230,293
Mean length (bp)	1,151	2,000	1,500	1,726
# of *G. affinis* BLASTP Hits	---	20,511	19,904	18,880

[1]http://www.ncbi.nlm.nih.gov/genome/annotation_euk/Poecilia_reticulata/100/

[2]http://www.ncbi.nlm.nih.gov/genome/annotation_euk/Xiphophorus_maculatus/101/

[3]http://www.ncbi.nlm.nih.gov/genome/annotation_euk/Oryzias_latipes/101/

Note: Table shows features of the protein encoding genes of *G. affinis* and their comparisons to those of closely related organisms (*Oryzias latipes*, *Poecilia reticulata*, and *Xiphophorus maculatus*). Data on features of the protein encoding genes of closely related organisms were obtained from NCBI (Links are below table). Results from validation of *G. affinis* predicted genes by BLAST analyses are also included.

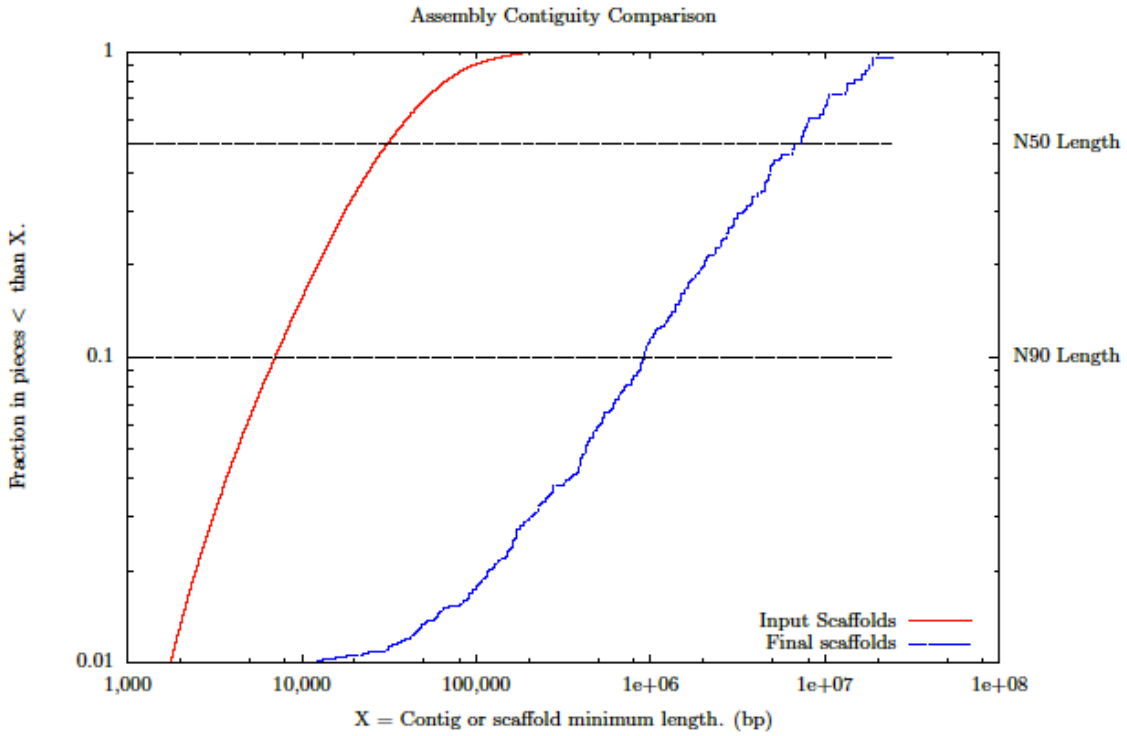
Table 4.3. Summary of repetitive elements in the genome of *Gambusia affinis*

Classification	Number of copies	Assembly coverage
DNA Transposons	318331	9.36138
Academ	600	0.02722
CMC-Chapaev-3	310	0.01579
CMC-EnSpm	825	0.01427
Kolobok-T2	607	0.02563
Merlin	1144	0.0303
MULE-MuDR	898	0.0329
PIF-Harbinger	9897	0.45702
PIF-ISL2EU	3559	0.16284
PiggyBac	7996	0.3492
TcMariner	183842	4.78829
hAT	62452	2.03977
Helitron	4743	0.19566
MITE	21578	0.56671
Unclassified DNA	19880	0.65578
LTR Retrotransposons	12602	0.37873
Gypsy	1471	0.07761
Ngaro	10160	0.26536
Pao	287	0.00657
ERV1	454	0.01758
Unclassified LTR	230	0.01161
LINE Retrotransposons	50048	1.40137
Dong-R4	195	0.00787
I	341	0.00961
LINE1	670	0.01869
LINE2	26519	0.72455
Penelope	770	0.01738
Proto2	136	0.00544
RTE	7158	0.21687
Rex-Babar	13841	0.38951
Unclassified LINE	418	0.01145
SINE Retrotransposons	16609	0.4273
MIR	2100	0.04935
tRNA	14289	0.3677
Unclassified SINE	220	0.01025
Unknown	198564	6.23
Interspersed Repeats	596154	17.79878
Low complexity regions	33073	0.25
Satellites	4914	0.23
Simple repeats	219965	1.43
TOTAL Repeats	854106	19.70878

Definitions

Low complexity regions	Rich regions: A-rich, GA-rich, C-rich…
Simple repeats	Duplications of simple sets of DNA bases (1-5bp) such as A, CA, CGG
Satellites	Duplications of more complex sequences (100-200 bases)
Interspersed repeats	Transposable elements: DNA transposons, LTR, LINE, SINE Retrotransposons Unknowns are usually included in interspersed

Figure 4.1 Contiguity Comparisons



Note: A comparison of the contiguity of the input assembly and the _nal HiRise sca_olds. Each curve shows the fraction of the total length of the assembly present in sca_olds of a given length or smaller. The fraction of the assembly is indicated on the Y-axis and the sca_old length in basepairs is given on the X-axis. The two dashed lines mark the N50 and N90 lengths of each assembly. This plot excludes sca_olds less than 1 kb.

Figure 4.2 Synteny Analyses

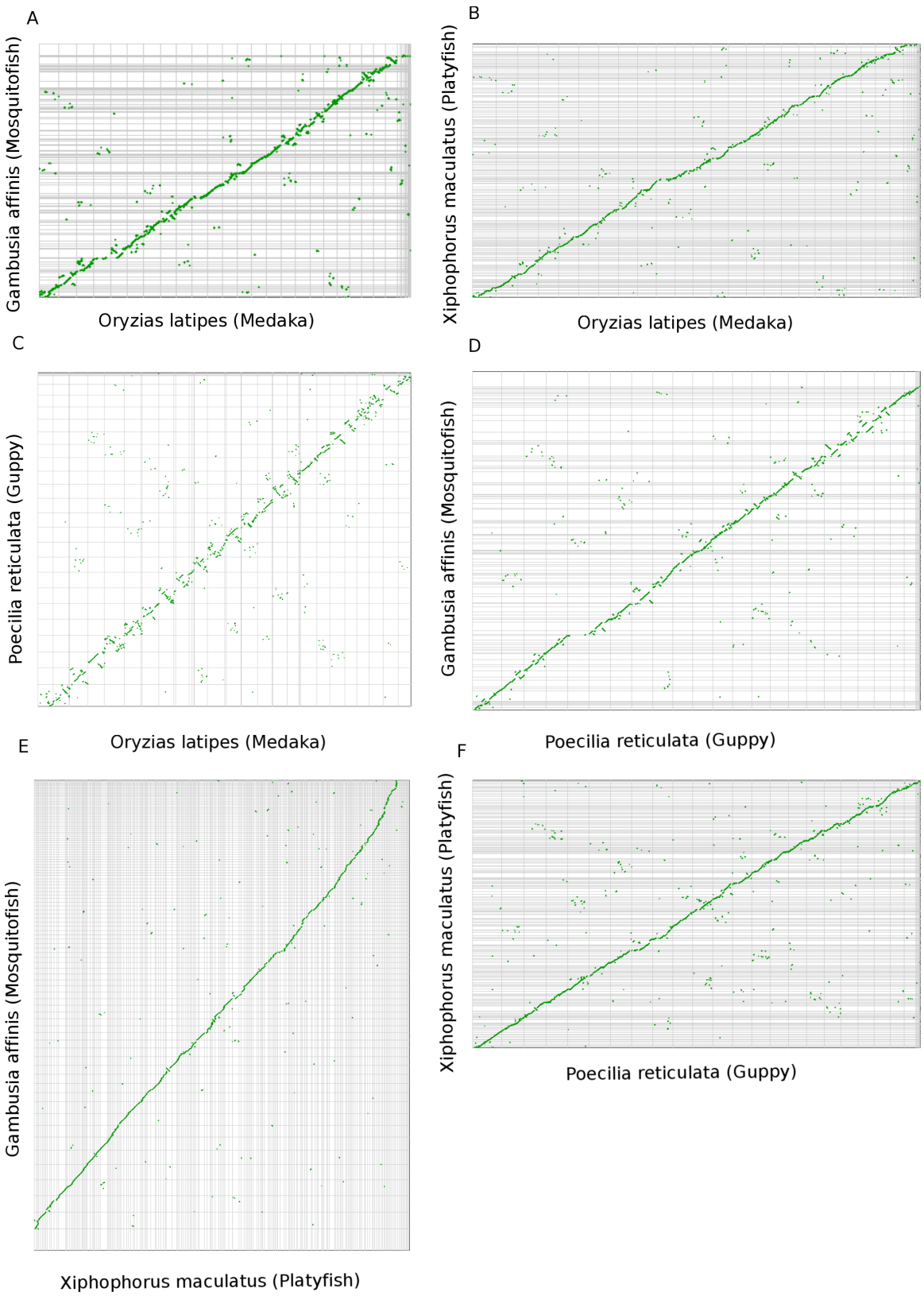


Figure 4.2 shows the analyses of synteny between *Gambusia affinis* and three other fish species, two Poeciliidae and Medaka. Figures were generated using CoGe Synmap. Panel A shows *G. affinis* mapped to *Oryzias latip*es (medaka). Panel B shows *X. maculatus* mapped to *O latipes*.. Panel C shows *P. reticulata* mapped to *O. latipes.* Panel D. shows *G. affinis* plotted against the guppy *Poecilia reticulata*. Panel E shows *G. affinis* mapped to the platyfish *Xiphophorus maculatus*. Panel F shows *X. maculatus* mapped against *P. reticulata*. All analyses were limited to scaffolds larger that were 1MB or larger.

Appendix C

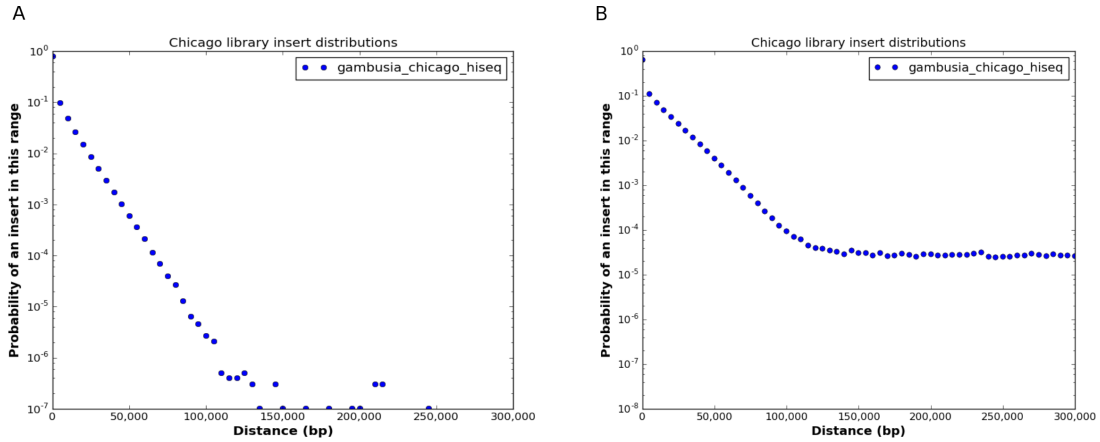


Figure S.4.1. Panel A shows the distribution of probabilities of inserts in the Meraculous assembly. Panel B shows the distribution of probabilities of inserts in the HiRise assembly.

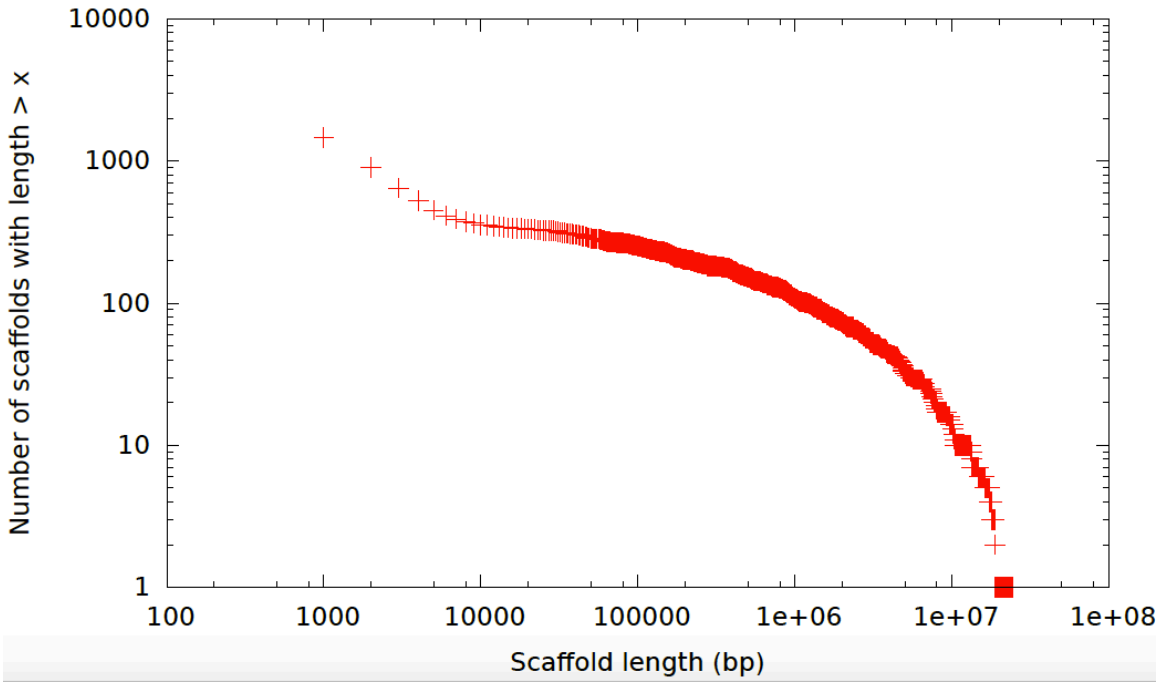


Figure S.4.2. Depicts the distribution of scaffold lengths in bp in the final HiRise genome assembly.

CHAPTER 5: ASSESSING THE GENETIC DIVERSITY OF THE INVASIVE WESTERN MOSQUITOFISH *GAMBUSIA AFFINIS*[4]

[4] Troendle NJ Glenn T and Mauricio R. To be submitted to *Biological Invasions*

Abstract

When a species is introduced it is unlikely that the new population will contain all of the genetic diversity found in the native range. This should prove to be a disadvantage when invasive species are subjected to new evolutionary pressures, yet invasive species still thrive in most regions of the world, a phenomenon known as the 'genetic paradox.' This loss of genetic diversity is likely mediated either through large introductions or multiple introductions. We used RadSeq data to produce a suite of 4405 SNPs for 12 populations of the invasive Western Mosquitofish *Gambusia affinis* along its proposed invasion route to investigate the amount of genetic diversity in the invasive range of the species. We found reductions in genetic diversity in some locations of the invasive range though not to the degree that might be expected in the case of serial introductions. We also found that in some parts of the invasive range diversity as high or higher than in the native range including the source population for the invasion to Asia, which indicates that there have been multiple introductions into these regions which has caused this increase in diversity.

Introduction

Invasive species are responsible for the erosion of biodiversity and habitat destruction around the world (Clavero, 2005; Gureveitch & Padilla, 2004; Mooney & Cleland 2001; Dextrase & Mandrake, 2006). However, despite doing such damage invasive species do provide opportunities to study evolution as it occurs as invasive species face novel selection pressures and are often introduced widely into various habitats making them natural experiments in evolution (Sakai *et al.* 2001; Huey *et al.* 2005). As a result, invasive species are widely studied by ecologists and evolutionary biologists.

Many invasive species are the result of human introductions (McNeely, 2001). When humans introduce species, whether intentionally or not it is unlikely that the introduced population captures all of the genetic diversity in the source population (Roman and Darling, 2007). Despite this expected decrease in genetic diversity many invasive species still manage to flourish (Tsutsui *et al.* 2000). This is often called the "genetic paradox" because we expect populations with lower diversity to be less fit but invasive populations often thrive (Frankham, 2005). While some species do have decreased diversity in their invasive ranges this tend is not as universal as it was originally thought to be (Roman and Darling, 2007). Kolbe *et al.* (2004), found that in anolis lizards in Cuba genetic diversity was actually higher than in the native range. Genton *et al.* (2005), similarly found that there was no decrease in diversity in the invasive range of *Ambrosia artemisiifolia*. Similar patterns have been found in several invasive species and have been attributed to multiple introductions (Dlugosch and Parker, 2008). By having multiple introductions into a location, recombination and introduction of new alleles can lead to levels of genetic diversity as high or sometimes higher than in the native range (Hufbauer, 2008).

This study focuses on the western mosquitofish, *Gambusia affinis,* which is native to the southeastern United States, but was introduced throughout the world in the early 20th century as a means of mosquito control in areas where malaria and yellow fever were common (Krumholz 1948; Pyke 2008). While their efficacy is a subject of much debate, *G. affinis* and its sister species *G. holbrooki* have become extremely widespread and today are present in over 50 countries (Pyke, 2008; Lowe *et al.* 2000). Historical records exist which suggest that *G. affinis* has undergone serial introductions from Texas, to Hawaii, and then from Hawaii to Taiwan and the Philippines. Taiwan and the Philippines they were then introduced to Japan and China ((Jordan, 1927; Yan *et al.* 2001; Xie *et al.* 2010; Seale 1917; Koya *et al.* 1998; Pan *et al.* 1980). Due to the nature of these introductions genetic diversity would be expected to be lower in the invasive range and each subsequent introduction should have a decrease in diversity (Purcell *et al.* 2012).

Lee (2014), used microsatellites and mitochondrial sequencing in an attempt to determine the validity of the historical records. They found evidence that supported the idea of serial introductions resulting in a decrease in genetic diversity along the invasion route. Lee (2014), found no evidence of a genetic bottleneck dispite reduced genetic diversity in the invasive range which would indicate that either the introductions were made with very large numbers of fish or there had been multiple introductions.

The purpose of this study is to answer two questions. First, is there a decrease in the amount of genetic diversity between the native and invasive ranges of *G. affinis* as is typically expected in cases of serial introductions? And second, is there rescuing of genetic diversity in the invasive range indicative of either large or multiple introductions?

Methods

Collections

I selected one hundred twelve *Gambusia affinis* from the collections made in Lee (2014), and used the DNA samples that they extracted. These collections covered the native range of the species as well as the route of invasion to Asia as proposed by the historical records (Seale, 1905; Seale, 1917; Jordan, 1927). From these collections I selected 112 fish choosing 8 fish from 12 populations along the invasion route and 16 from the proposed putative source population (Table 5.1). I selected these localities based on the proposed invasive route, and the quality and quantity of DNA available.

RadSeq Library Preparation

I performed RADseq according to the 3RAD method of Graham *et al.* (2015). I digested samples in a 15ul reaction, comprised of 100ng of DNA, 10 units of XbaI, 10 units of EcoRI-HF, 10 units of NheI-HF (New England Biolabs, Ipswich, MA, USA), 1x Cutsmart Buffer (New England Biolabs, Ipswich, MA, USA), 1µl of forward and reverse adapter at 5µM. Each sample had a unique adapter combination. I incubated samples for 1 hr at 37˚C after which I immediately added 100,000 units of T4 DNA ligase (New England Biolabs, Ipswich, MA, USA), 0.5µl of 10x ligase buffer (New England Biolabs, Ipswich, MA, USA), and 1.5µl of 10mM rATP (Promega, Madison, WI, USA) in 5µl to each sample. I then incubated samples at 22˚C for 20 mins and 37˚C for 10 mins for 2 cycles and then 80˚C for 20 mins. I pooled 10uL of each sample post-ligation product and cleaned with diluted (Rohland & Reich 2012) Sera-mag Speedbeads (Fisher, Pittsburgh, PA, USA) in a 1:1 ratio, washed twice with 80% EtOH and resuspended in

50ul TLE (10 mM Tris pH8, 0.1 mM EDTA).

I used the post-ligation pool as the template for triplicate PCR using dual-indexed primers to uniquely tag each sample (Glenn *et al.* 2016). In 25µl reactions, 10µl of DNA was combined with 2.5ul of 5µM i5 primer, 2.5ul of 5µM i7 primer, 0.75 µl of 10mM dNTPs, 5ul of 5x KAPA HiFi Fidelity Buffer, and 0.5 unit of KAPA HiFi Hotstart DNA Polymerase (KAPA Biosciences, Boston, MA, USA). I incubated these at 95˚C for 2min., then 12 cycles of 98˚C for 20sec, 60˚C for 30 sec, 72˚C for 1 minute; followed by 72˚C for 5 mins.

I then pooled PCR samples and cleaned them using a 1:1 DNA to diluted Speedbead ratio, washed twice with 80% EtOH and resuspended in 25µl TLE. I quantified and combined the pool for sequencing and size selection. I performed size selection on the pool of libraries with PippinPrep (Sage Science, Beverly, MA, USA) using a CDF1510 agarose gel cassette and size range set on tight at 550bp. After size selection, I recovered DNA from the collection well and cleaned it using a 1:1 DNA to diluted Speedbead ratio, I washed twice with 80% EtOH and resuspended in 30ul TLE. After cleanup, I performed PCR with the P7 and P5 primers. The reaction included 20µl of DNA, 5µl of 5µM of each primer, 25uL KAPA HiFi Hotstart ReadyMix (KAPA Biosciences, Boston, MA, USA) in a 50µl reaction. I incubated samples at 95˚C for 45 seconds; 6 cycles of 98˚C for 15 sec, 60˚C for 30 sec, 72˚C for 1 minute; and 72˚C for 5 min. I then quantified and combined my PCR for sequencing on an Illumina NextSeq High Output v2 150 cycle kit to obtain paired-end 75 reads.

Stacks

Of my 112 original samples 86 had sufficient reads to be analyzed via the Stacks pipeline (Catchen, 2011; Catchen, 2013). I used the program BWA (Li and Durbin, 2009), with the

backtrack algorithm to map my 86 RadSeq samples to the *Gambusia affinis* genome described in chapter 4, (unpublished data). I then used the pstacks pipeline to search for single nucleotide polymorphisms (SNPs) in these 86 samples. I extracted these SNPs using the Populations portion of the Stacks pipeline with the requirements that 80% of individuals within at least one population have a particular locus for it to be included. I set minimum stack depth at 32 with a minimum minor allele frequency of 0.1. I specified a maximum heterozygosity of 1 and limited the data analysis to the first SNP at each locus. I also used the Populations portion of the pipeline to generate basic population genetics statistics along the invasion route, based on my SNPs.

Neighbor Joining Tree

I constructed 1352 population Neighbor Joining trees with Geneious (Kearse *et al.* 2012), using three distance models, (Tamura-Nei, HKY and Jukes Cantor). I generated consensus trees by sorting topologies and included support values at each node. I compared the three distance models in order to confirm the consistency of my results. As all three distance models produced identical optimal topologies I chose to display the Tamura-Nei tree since this is the most common distance model used. I considered trees with identical topologies but different branch supports to be identical.

Structure

In order to determine the population structure among my populations of interest, I used the program STRUCTURE (Pritchard *et al.* 2000). I used my dataset of 4405 SNPs with admixture and assumed no correlation between SNPs. I performed three iterations for K values

1-13 with random seeds. Each run consisted of a 50,000 burnin period which was discarded and an additional 100,000 generations. I then used STRUCTURE HARVESTER to identify the optimal K value based upon delta K (Earl and VonHoldt 2011). After that I used CLUMPP to determine the most likely cluster membership coefficients for the optimal K value using the greedy algorithm (Jakobsson and Rosenberg 2007). Finally I visualized my data using Excel and Inkscape (Harrington *et al.* 2004).

Measures of Diversity

To estimate genetic diversity I performed a Global Hardy-Weinberg test in Genepop (Rousset, 2008) testing for heterozygote excess in each population and comparing the native and invasive regions. I also used GenAlEx (Peakall and Smouse, 2006; Peakall and Smouse, 2012), to calculate how many private alleles each individual had. I then performed a comparison of means test, using the Dunnett's method in JMP (SAS, 2012), comparing each population to the source population, Pine Gully. I also performed a Wilcoxon/Kruskal-Wallis test in JMP to determine if there were a significant difference in the number of private alleles in each range. I also calculated the percentage of loci that were polymorphic for each population in GenAlEx (Peakall and Smouse, 2006; Peakall and Smouse, 2012). Finally, I used GenAlEx to perform an Analysis of Molecular Variance to determine the amount of variance distributed between the native and invasive ranges and between and within populations (Table S.5.1).

Results

Overall, I found that genetic diversity did not decrease as much as expected in the invasive range. While some measures of diversity did show decreases others showed no

significant changes when looking at the invasive range as a whole. I found that heterozygosity did not significantly differ between the ranges (Hardy-Weinberg Global test of heterozygote excess; $p = 1$, $SE = 0$). I did, however find that the number of private alleles per individual was lower in the invasive range than in the native range (Kruskal-Wallis 2 sample test; $S = 1620$, $Z = 5.95$, $p < 0.0001$). The percentage of loci that were polymorphic also tended to be lower for most populations in the invasive range (Table 5.2). The STRUCTURE analysis shows several of the populations in the invasive range showing cluster identity to only two or three clusters, which could indicate fixation at several of loci due to a reduction in the number of alleles as a result of a founder event.

While some metrics indicate a decrease in genetic diversity in the invasive range, this is not a universal trend across populations. Each population along the invasive route has different levels of diversity. In some of these locations, I observed a decrease in diversity. For example, the population at Xishuangbanna Tropical Botanical Garden has significantly fewer private alleles than the source population, Pine Gully (Comparison of means test PG mean private alleles = 277.06, XTBG mean private alleles = 1.86, Abs(Dif)-LSD = 81.29, $p = 0.0011$, Table 5.4). This population also has significantly lower heterozygosity than the source population (Hardy-Weinberg Global Test; $p < 0.0001$, $SE = 0$, Table 5.3).

However, in one locality, Shanghai Ocean University, in China, I found a significant increase in the number of private alleles (Comparison of means test PG mean private alleles = 277.06, Shanghai mean private alleles = 1854.2, Abs(Dif)-LSD = 1358, $p < 0.0001$, Table 5.4). Shanghai Ocean University also had no significant decrease in heterozygosity from the native range (SH H_o=0.3847 and PG H_o=0.3226; Hardy-Weinberg Global Test; $p=0.7907$, $SE=0.0018$, Table 5.3). The population in Shanghai has the highest proportion of loci that are polymorphic

in the invasive range (Table 5.2). I found that the JIJI locality in Taiwan has no significant difference in either heterozygosity or number of private alleles from the source population (JIJI H_o = 0.2684 and PG H_o = 0.3226; Hardy-Weinberg Global Test; p = 0.994, SE = 0.0003, Table 5.3 Comparison of means test PG mean private alleles = 277.06, JIJI mean private alleles=52, Abs(Dif)-LSD = 5.819, p = 0.2919, Table 5.4).

I generated 1352 trees using three distance metrics and chose the tree with the optimal topology. Each of the three distance metrics used in my phylogenetic analyses yielded identical optimal topologies but with different branch supports. The tree generated by the Tamura-Nei distance measure was chosen for display (Figure 1). The tree that I generated supports the historical record of the invasion route. Pine Gully is located at the base of the tree, as we would expect of the putative source population. I then have clusters of populations that are indicative of serial introductions, with populations within a locality grouping close to one another but also grouping close to their potential source locality. The Chinese populations appear to be the most recently established populations according to my tree with XTBG and Lover's Lake being the last two to diverge. The exception to this is the Shanghai Ocean University location, which is grouped with the Hawaiian locality and the Yilian University locality in Taiwan. Finally, the Johnson Creek locality from the native range grouped with the populations from the Philippines.

I used STRUCTURE HARVESTER and found that the optimal number of clusters was K = 10 (Figure S.1). My analysis of population structure showed evidence for multiple introductions into the area of the Shanghai locality (Figure 2). Shanghai shared cluster identity with the Yilian University locality in Taiwan as well as clusters present in the Midori River locality in Japan and Guagua in the Philippines. Shanghai also shared much of its cluster identity with the South Lake locality in China, which is the closest of my populations

geographically. The XTBG population, which is located in a botanical garden in Yunnan province, had very low diversity grouping to a single cluster for all but one individual. Both clusters that the XTBG locality groups with are also found in the nearest locality Lover's Lake (Figure S.2).

Discussion

Lee (2014) used microsatellites and mitochondrial data to measure diversity in the native and invasive range of *G. affinis*. He found that there was significantly lower diversity in the invasive range as would be expected in the case of serial introductions. When you have a scenario with serial introductions, that is, one introduced population becomes the source for new introductions I expect genetic diversity to decrease with each step in the route. This is because each time you subsample a population as you do when using them for an introduction it is unlikely that all of the genetic diversity is captured (Nei *et al.* 1975). This can be overcome however, if you introduce the species multiple times from different sources (Dlugosh and Parker, 2008). My analysis of the genetic diversity along the invasion route shows that overall there is a decrease in diversity in the invasive range but only in some populations and not for all measures of diversity. While many populations have fewer private alleles than the native range, this is not universal as the population in Shanghai actually has significantly more private alleles than are present in the native range. Heterozygosity in the Shanghai population is also not significantly different from the source population Pine Gully. I also found that the percentage of loci that were polymorphic was higher in the Shanghai population than in the rest of the invasive range. So, while we do see reductions in genetic diversity in parts of the invasive range this does not hold true for the Shanghai population. This indicates that there have been reductions in diversity

in these regions though it has been mediated in Shanghai by multiple introductions or through very large introductions. This combined with the high level of heterozygosity suggests that this locality may have recovered from the founder's events that often reduce genetic diversity (Barrett and Husband, 1990). While there are other populations that do not display decreased heterozygosity, none of them show an increase in the number of private alleles like the Shanghai population does.

We know from the historical records of introduction that there have been multiple introductions to China. The first introduction came from Taiwan (Yan *et al.* 2001) in 1924 but there is no mention of how many fish were introduced or what area they were released in. The second introduction was from the Philippines into the region around Shanghai in 1927 and the third was also from the Philippines into Guangzhou in 1960. My STRUCTURE analysis supports the theory that there have been several introductions into China. While very little is known about the introduction from Taiwan, my STRUCTURE and phylogenetic analysis both indicate that there was likely at least one introduction from the Yilian University region of Taiwan into the area around Shanghai which is indicated by their cluster identity. My analyses also validate the claim of an introduction from the Philippines to Shanghai. The Shanghai and South Lake localities both identify with clusters that are present at some frequency in the Philippines, which is indicative of a founder's event, when a small group of individuals is introduced to a new area reducing genetic diversity. My analysis can verify at least two introductions into the region, however, because Shanghai is such a major port and given all of our measures of genetic diversity are as high or higher than in the putative source population, it is possible that there have been more than these two introductions.

I found that while the invasive range did have lower diversity in some locations this was not universal and some locations had genetic diversity as high or higher than the native range. Based upon the relatively high level of genetic diversity in some locations in the invasive range when compared to the native range, it is likely that introductions have either been extremely large, which is not always the case according to the records, or there have been multiple introductions to these regions (Dlugosch and Parker 2008). This is actually a common phenomenon and is known as the 'genetic paradox' (Allendorf and Lundquist 2003).

References

Allendorf, F., L. Lundquist. 2003. Introduction: Population biology, evolution and control of invasive species. Conservation Biology **17**: 24-30.

Barrett, S.C., B.C. Husband, A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir. 1990. The genetics of plant migration and colonization. Plant Population Genetics, Breeding, and Genetic Resources 254-277.

Catchen, J., P. Hohenlohe, S. Bassham, A. Amores, and W. Cresko. 2013. Stacks: An analysis tool set for population genomics. Molecular Ecology **22**: 3124-3140.

Catchen, J., A. Amores, P. Hohenlohe, W. Cresko, and J. Postlethwait. 2011. Stacks: Building and genotyping loci *de novo* from short-read sequences. G3: Genes, Genomes, Genetics **1**: 171-182.

Clavero, M. and E. Garcia-Berthou. 2005. Invasive species are a leading cause of animal extinctions. Trends in Ecology and Evolution **20**: 110-110.

Dextrase, A.J., and N.E. Mandrak. 2006. Impacts of alien invasive species on freshwater fauna at risk in Canada. Biological Invasions **8**: 13-24.

Dlugosch K.M., and I.M. Parker. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology **17**: 431-449.

Earl D.A., and B.M. VonHoldt. 2011 STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservaton Genetic Resources **4**: 359–361.

Frankham, R. 2005. Resolving the genetic paradox in invasive species. Heredity **24**: 385-385.

Genton, B.J., A. Shykoff, and T. Giraud. 2005. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia,* as a result of multiple sources of introduction. Molecular Ecology **14**: 4275-4285.

Glenn, T.C., R. Nilsen, T.J. Kieran, J.W. Finger, T.W. Pierson, K.E. Bentley, S. Hoffberg, S. Louha, F.J. Garcia-De-Leon, M.A. del Rio Portilla, and K. Reed. 2016. Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). bioRxiv, p.049114.

Graham, C.F., T.C. Glenn, A. McArthur, D.R. Boreham, T. Kieran, S. Lance, R.G. Manzon, J.A. Martino, T. Pierson, S.M. Rogers, J. Wilson, C.M. Somers. 2015. Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). Molecular Ecology Resources **15**: 1304-1315.

Gurevitch, J. and D.K. Padilla. 2004. Are invasive species a major cause of extinctions? Trends in Ecology and Evolution **19**: 470-474.

Huey, R.B., G.W. Gilchrist, and A.P. Hendry. 2005. Using invasive species to study evolution. Pp. 138-164 *in*: D.F. Sax, J.J. Stachowicz, and S.D. Gaines (editors). Species Invasions: Insights to Ecology, Evolution and Biogeography. Sinauer Associates, Sunderland, MA.

Hufbauer, R. 2008. Biological invasions: Paradox lost and paradise gained. Current Biology **18**: 246-247.

Jakobsson, M., and N.A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics **23**: 1801-1806.

Jordan, D. 1927. The mosquito fish (*Gambusia*) and its relation to malaria. Report to Board Regents Smithsonian Institute **1926**: 361-368.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Mentjies, and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28**: 1647-1649.

Kolbe, J.J., R.E. Glor, L.R. Schettino, A.C. Lara, A. Larson, and J.B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature **431**: 177-181.

Koya, Y., T. Itazu, and M. Inoue. 1998. Annual reproductive cycle based on histological changes in the ovary of the female mosquitofish, *Gambusia affinis,* in central Japan. Ichthyology Research **45**: 241-248.

Krumholz, L. 1948. Reproduction in the western mosquitofish, *Gambusia affinis affinis* (Baird & Girard), and its use in mosquito control. Ecological Monographs **18**: 1-43.

Lee, J. B. 2014. Tracing invasions by comparing native and introduced populations using empirical and simulated data. Ph.D. dissertation, University of Georgia, Athens, GA, USA.

Li H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. **25**: 1754-1760.

Lindholm, A.K., F. Breden, H.J. Alexander, W.K. Chan, S.G. Thakurta, and R. Brooks. 2005. Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. Molecular Ecology **14**: 3671-3682.

Lowe S., B.M. Boudjelas, S.M. De Poorter. 2000. 100 of the world's worst invasive alien species. The Invasive Species Specialist Group **11**: 1-13.

McNeely, J.A. 2001. The Great Reshuffling: Human Dimensions of Invasive Alien Species. IUCN, Cambridge UK.

Mooney, H.A. and E.E. Cleland. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Science, USA **98**: 5446-5451.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution **29**: 1-10.

Pan, J., B. Su, and W. Zheng. 1980. Biological characteristics of *Gambusia affinis* and the prospects for its use in for mosquito control. Journal South China Normal University (Natural Sci) 117-138.

Peakall, R., and P.E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes. **6**: 288-295.

Peakall, R., and P.E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics **28**: 2537-2539.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics **155**: 945-959.

Purcell K.M., and L.N. Stockwell. 2012. Evaluation of the introduction history and genetic diversity of a serially introduced fish population in New Zealand. Biological Invasions **14**: 2057–2065.

Pyke, G. H. 2008. Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. Annual Review of Ecology, Evolution, and Systematics **39**: 171-191.

Roman, J., and J.A. Darling. 2007. Paradox lost: Genetic diversity and the success of aquatic invasions. Trends in Ecology and Evolution **22**: 454-464.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources **8**: 103-106.

Sakai A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, With A. Kimberly, S. Baughman, R.J. Cabin, J.E. Cohen, N.C. Ellstrand, D.E. McCauley, P. O'Neil, I.M. Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annual Review of Ecological Systems **32**: 305-332.

Seale, A. 1905. Report of Mr. Alvin Seale of the United States Fish Commission, on the introduction of top-minnows to Hawaii from Galveston, Texas. Hawaiian Forester Agriculture **2**: 364-367.

Seale, A. 1917. The mosquito fish, *Gambusia affinis* (Baird and Girrard), in the Philippine Islands. Philippine Journal of Science **12**: 177-187.

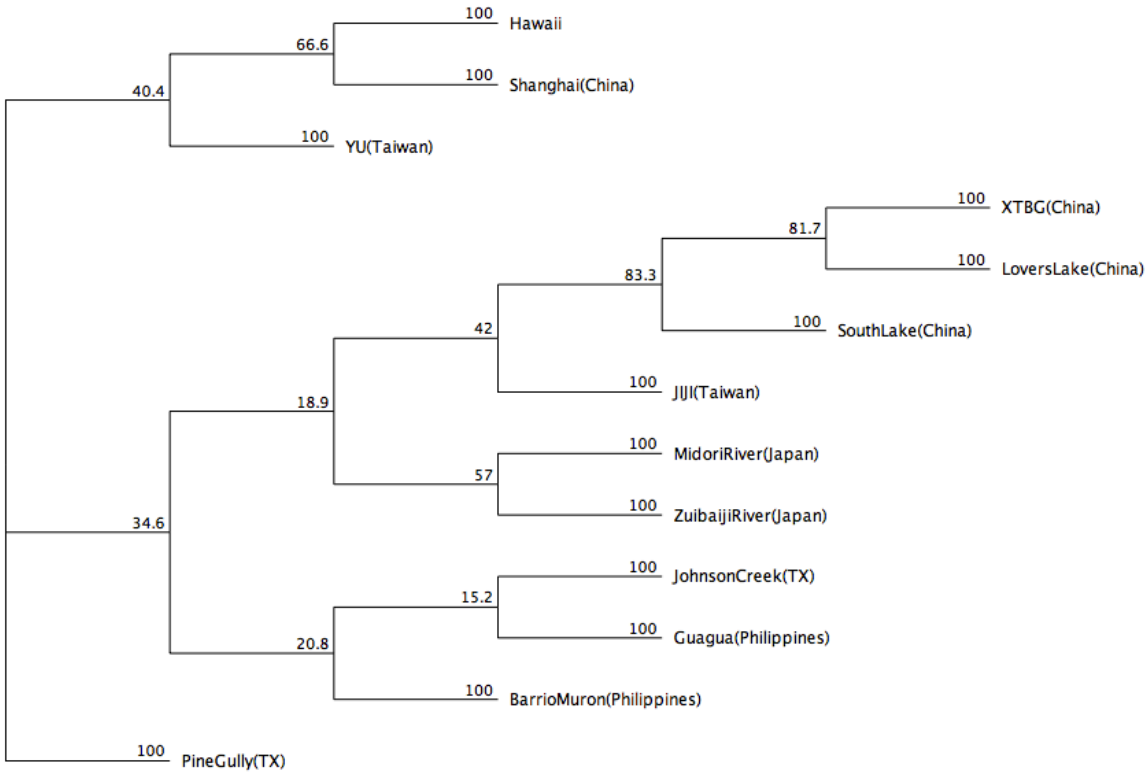
Tsutsui, N.D., A.V. Suarez, D.A. Holway, and T.J. Case. 2002. Reduced genetic variation and the success of an invasive species. Proceedings of the National Academy of Science USA **97**: 5948-5953

Xie Y.P., Z.Q. Fang, L.P. Hou, and G.G. Ying. 2010. Altered development and reproduction in western mosquitofish (*Gambusia affinis*) found in the Hanxi River, southern China. Environmental Toxicology and Chemistry **29**: 2607–2615.

Yan, X., L. Zhenyu, W. Greg, and L. Dianmo. 2001. Invasive species in China-an overview. Biodiversity Conservation **10**: 1317-1341

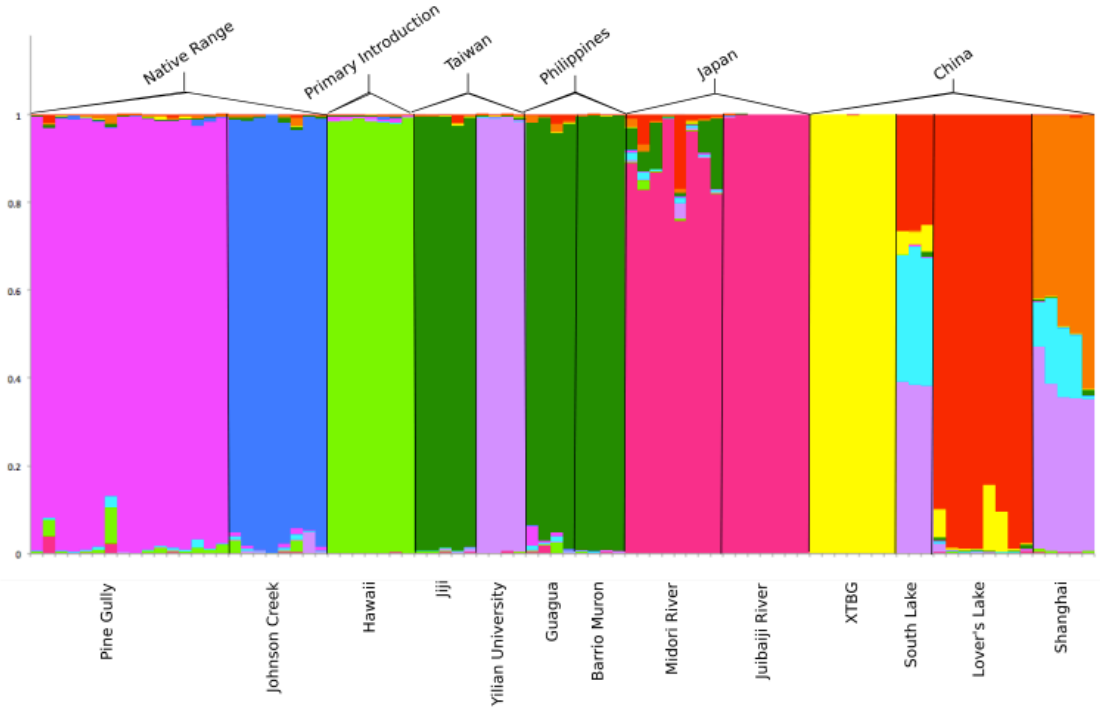
Figures and Tables

Figure 5.1: Tamura-Nei Neighbor Joining Tree



Note: Neighbor Joining Tree constructed from 4405 SNP loci, using Tamura-Nei distance measurements. Branch labels are consensus percentage support values based on optimal topology.

Figure 5.2: Structure Plot for K=10



Note: Structure plot based on 4405 SNPs. K=10 as indicated by structure harvester to be the optimal K value. Collection localities are shown below plot with region of collection above.

Table 5.1: Collection Sites

Label	Locality Name	Location	Region	N	Lat.	Long.
PG	Pine Gully	Texas	Native	16	29.59	-95
JC	Johnson Creek	Texas	Native	8	30.15	-99.34
HI	Kualoa Regional Park	Hawaii	Invasive	8	21.51	-157.84
YU	Yilan University	Taiwan	Invasive	8	24.75	121.74
JIJI	Jiji	Taiwan	Invasive	8	23.83	120.8
GU	Guagua	Philippines	Invasive	8	14.96	120.64
BM	Barrio Muron	Philippines	Invasive	8	14.67	120.98
MR	Midori River	Japan	Invasive	8	32.75	130.7
ZR	Zuibaiji River	Japan	Invasive	8	33.59	130.25
SH	Shanghai Ocean University	China	Invasive	8	30.88	121.9
SL	South Lake	China	Invasive	8	30.47	114.38
LL	Lover's Lake	China	Invasive	8	23.14	113.35
XG	Xishuangbanna Tropical Botanical Garden	China	Invasive	8	21.93	101.26

Note: This table shows the location of the 13 populations used for this study. Region=Whether the population was sampled form the native or invasive range; N=number of individuals sampled; Lat and Long give the coordinates of the collection site.

Table 5.2: Diversity Measures

Locality Name	Location	Region	H_O	H_E	% Polymorphic Loci
Pine Gully	Texas	Native	0.3226	0.3314	38.43
Johnson Creek	Texas	Native	0.2243	0.2122	18.43
Kualoa Regional Park	Hawaii	Invasive	0.2645	0.259	15.05
Yilan University	Taiwan	Invasive	0.2373	0.2057	9.04
Jiji	Taiwan	Invasive	0.2684	0.2496	17.68
Guagua	Philippines	Invasive	0.2872	0.261	11.37
Barrio Muron	Philippines	Invasive	0.2553	0.2386	12.35
Midori River	Japan	Invasive	0.2618	0.2513	11.94
Zuibaiji River	Japan	Invasive	0.2453	0.2176	12.12
Shanghai Ocean University	China	Invasive	0.3847	0.3421	59.48
South Lake	China	Invasive	0.2308	0.1928	9.31
Lover's Lake	China	Invasive	0.2726	0.2484	12.17
Xishuangbanna Tropical Botanical Garden	China	Invasive	0.1365	0.1118	5.54

Note: This table shows several diversity measurements for each population. H_o and H_E are observed and expected heterozygosities as calculated in the text using the Stacks Pipeline. % Polymorphic loci is the percentage of the 4405 SNP loci for which each population was not monomorphic.

Table 5.3 Global Hardy-Weinberg Test

Population	P value	Standard Error
Pine Gully	1	0
Johnson Creek	0.9779	0.0006
Hawaii	0.9996	0
JIJI	0.994	0.0003
YU	0.259	0.0017
G	0.9877	0.0004
BM	0.9998	0
MR	0.9695	0.0006
ZR	0.0281	0.0006
XTBG	<0.0001	0
SL	0.5566	0.0017
LL	0.0162	0.0005
SH	0.7907	0.0018

Note: This table displays differences in heterozygosities between each population and the source population (Pine Gully). P values and Standard errors calculated using Markov chain parameters (1000 Dememorizations; 100 Batches; 1000 Iterations/batch.

Table 5.4 Comparison of Means using Dunnett's Method

Population	Abs(Dif)-LSD	Mean	p-Value
Shanghai	1358	1854.2	<.0001
Pine Gully	-151	277.06	1
Johnson Creek	-47.7	139.5	0.2919
JIJI	5.819	52	0.0404
Yilian University	25.35	12.5	0.0208
Hawaii	71.86	11.29	0.0018
Zuibaiji River	72.29	10.86	0.0017
South Lake	2.835	5	0.046
Lovers Lake	88.39	3.38	0.0006
Bario Muron	34.85	3	0.0148
Midori River	89.39	2.38	0.0006
XTBG	81.29	1.86	0.0011
Guagua	-44.9	1	0.14

Note: Positive values of Abs(Dif)-LDS indicate means that are different from each other. p-Values are significant with a bonferoni corrected apha of 0.0042. Mean is the average number of private alleles for each population.

Appendix D

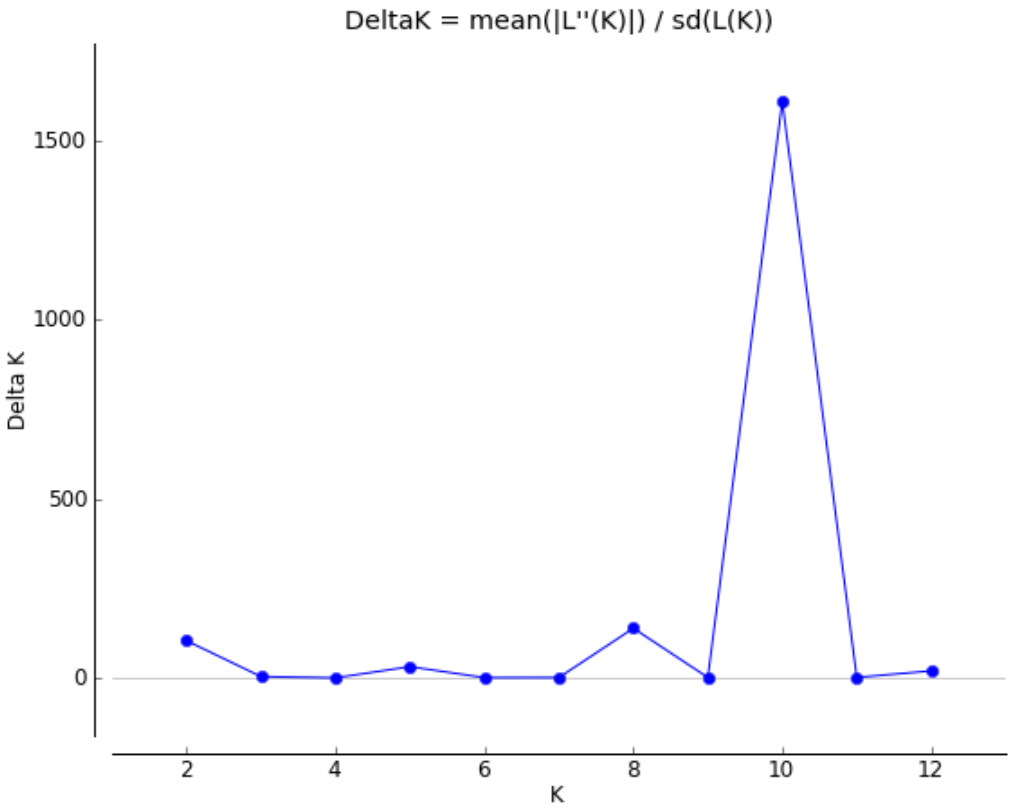


Figure S.5.1. Delta K graph generated in STRUCTURE HARVESTER. This shows K=10 as the optimal K value.

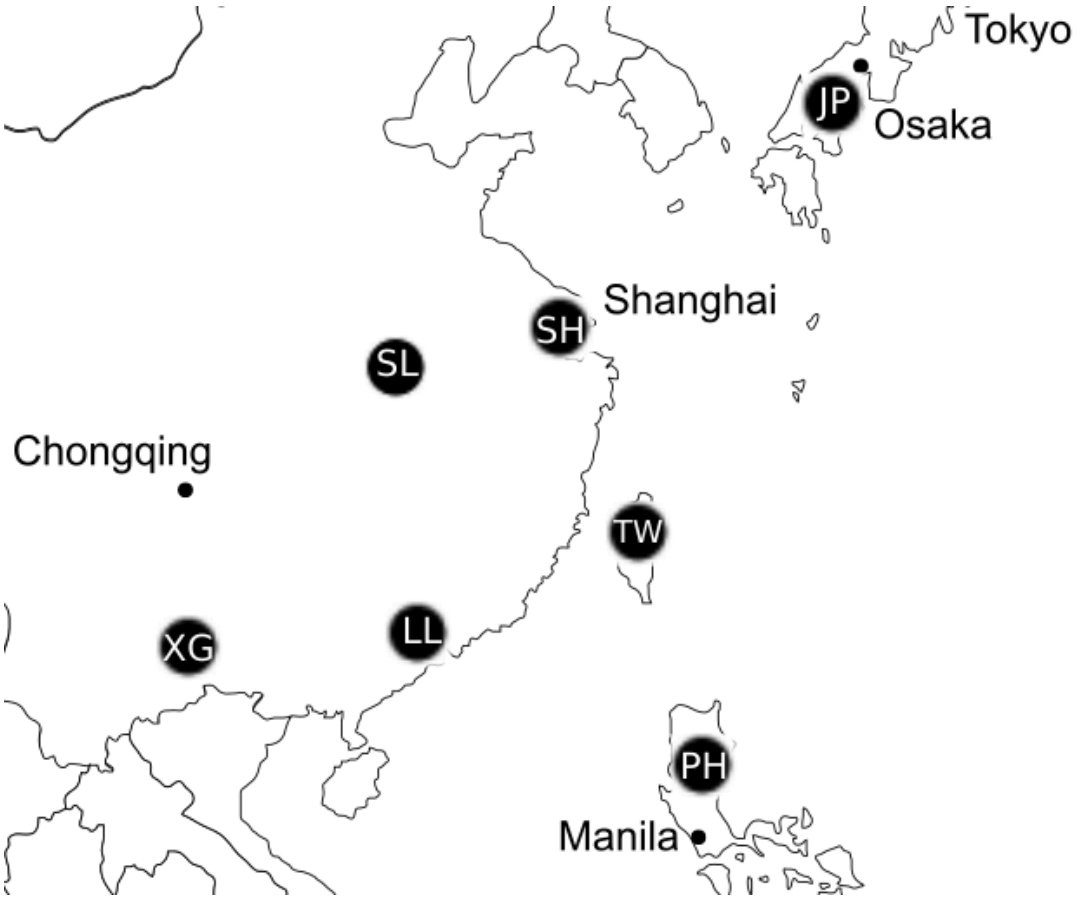


Figure S.5.2. Map of Asian collection sites in this study. XG= Xishuangbanna Tropical Botanical Garden, LL=Lover's Lake, SL=South Lake, SH=Shanghai, TW makes up both populations from Taiwan, PH consists of the two populations from the Philippines, JP= both Japanese populations.

Table S.5.1: Analysis of Molecular Variance (AMOVA)

Source	df	SS	MS	Est. Var.	%
Among Regions	1	14143.192	14143.192	84.422	12.96
Among Populations	11	50089.689	4553.608	354.416	54.40
Among Individuals	73	19567.311	268.045	55.345	8.49
Within Individuals	86	13532.500	157.355	157.355	24.15
Total	171	97332.692		651.538	100.00

Note: This table shows the results of the Analysis of Molecular Variance performed in GenAlEx on 4405 SNP loci. Regions are native and invasive ranges; %= the percentage of the variance accounted for by each effect.

CHAPTER 6: CONCLUSIONS

Invasive species have significant impacts on native species, environments and often economies (Elton, 1958; Charles & Dukes, 2007). There are several thousand invasive species in the United States alone and in addition to degrading the quality of environments, causing the decline of native species and introducing new diseases they do an estimated 1.4 trillion dollars in damage worldwide each year (Vitousek *et al.* 1996; Didham *et al.* 2005; Pimentel *et al.* 2001). In order to develop effective management strategies to control the spread and reduce the amount of damage done, we need to have a better understanding of invasive species (Sakai *et al.* 2001). Understanding the history of an invasion and where the species comes from can help to inform us on how to control it. However, since control is much more difficult once an invasive species becomes established we need to learn how to identify potential invasive species before they are introduced to new environments. This means we need to understand what makes an invasive species invasive. The most effective way to do this is to compare invasive and non-invasive species and identify the traits that make them different. This is best done in closely related species and has been done to some extent in plants (Rejmanek and Richardson, 1996), however vertebrate invasive species have been underrepresented in these types of studies.

In Chapter 2, I compared life history traits of 22 species of livebearing fishes from the family Poeciliidae. I specifically investigated the relationship between fecundity, reproductive investment, offspring size, and invasion status. I found that on average invasive species had more offspring and that those offspring tended to be smaller than their non-invasive relatives. I did not find any significant increase in the amount of resources invested in reproduction. This is

a common life history tradeoff that arises when the availability of resources remains constant but the energy from those resources is allocated differently depending on the specific life history strategy of that species. In this case since the proportion of resources allocated toward reproduction did not change a female fish must divide up those resources between each of her offspring. This means that she can either produce a small number of large offspring that require a lot of energy, which is what we see in the non-invasive species, or she can produce a large number of smaller offspring that require less energy individually (Roff, 2002). When I investigated each genus separately however, I found that this pattern was not universal. While, I observed this for the genera *Gambusia* and *Poecilia*, *Xiphophorus* and *Limia* did not show these changes. This illustrates the difficulty in producing a list of traits that we can say cause invasive behavior. The extreme amount of variation between species even within a family means that there is the potential for multiple invasive life history strategies.

In Chapter 3, I used the eleven invasive species from chapter 2 to study the changes in life history traits that occur when a species becomes invasive. Since invasive species are exposed to new environmental pressures they often evolve rapidly (Sakai *et al.* 2001). Because of this, measuring life history of invasive species only in the native range provides an incomplete picture of invasive life history. In order to better understand the link between life history and invasion I compared the life history of these eleven invasive species in their native and invasive ranges. I found that life history traits were significantly different in the invasive range. Offspring number, reproductive allotment and body size all increased in the invasive range while embryo size remained the same in the invasive range. This indicates that individuals in the invasive range have more resources overall allowing them to grow to greater size, devote more resources to reproduction by increasing the number of offspring without sacrificing embryo size.

It is important to note that the greater differences were observed in the genera *Limia* and *Xiphophorus*. These two genera did not show significant differences in life history when comparing invasive and non-invasive species in chapter 2, however they have significant changes when comparing their ranges. This indicates that for these two genera, invasive success may be driven by the flexibility of their life history traits and their ability to shift these traits when they become invasive. The genus *Poecilia,* however, showed very little change in the invasive range. Finally, I found that the invasive species from the genus *Gambusia* not only had significant higher fecundity, but that their fecundity increased in their invasive range without decreasing their offspring size further. This could be one of the reasons that the invasive *Gambusia* species are the most widespread freshwater fish in the world (Pyke, 2008)

In Chapter 4, we produced a high quality reference genome for the invasive western mosquitofish, *Gambusia affinis*. We used Illumina short read sequencing in conjunction with the new Chicago library HiRise assembly to increase contiguity and reduce fragmentation. The assembly produced has an N50 contig size of 31KB and an N50 scaffold size of 6.65MB at over 50X coverage. This reference genome will provide an excellent resource for the further study of the evolution of invasiveness as well as other aspects of Poeciliid research.

In Chapter 5, I used the reference genome described in chapter 4, along with RadSeq data to generate a panel of 4405 SNPs. I then used these SNPs to perform phylogenetic and statistical analyses in order to determine the degree of genetic diversity in *Gambusia affinis* in its invasive range. While some parts of the invasive range showed decreased genetic diversity this was not the case for the area around Shanghai, China. The amount of genetic diversity in the invasive range indicates that introductions have been large and that there have been multiple introductions

to some regions. I found evidence of at least two introductions to the area around Shanghai, China.

In conclusion, while invasive species have many negative impacts on their environments, native species and the economy, they do provide unique opportunities for studying evolution in real time (Sakai *et al.* 2001). By studying invasive species in their native and invasive ranges we can observe evolution, and since human-mediated introductions often occur on a contracted time scale and invasive species tend to evolve rapidly we are sometimes able to see these changes within decades rather than thousands of years. These types of studies will help us not only to understand and manage invasive species but will also increase our understanding of evolutionary principles.

References

Charles, H., and J.S. Dukes. 2007. Impacts of Invasive Species on Ecosystem Services. Pp. 217-237 *in:* W. Nentwick (editor). Biological Invasions. Springer, Berlin, Germany.

Didham, R.K., J.M. Tylianakis, M.A. Hutchinson, R.M. Ewers, and N.J. Gemmell. 2005. Are invasive species the drivers of ecological change? Trends in Ecology and Evolution **20**: 470-474.

Elton C.S. 1958. The ecology of invasions by animals and plants. Progress in Physical Geography **31**: 659-666.

Pimentel, D., S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. O'Connell, E. Wong, L. Russel, J. Zern, T. Aquino, and T. Tsomondo. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. Agriculture, Ecosystems and Environment **84**: 1-20.

Pyke, G. H. 2008. Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. Annual Review of Ecology, Evolution, and Systematics **39**: 171-191.

Rejmanek, M., and D.M. Richardson. 1996. What attributes make some plant species more invasive? Ecology **77**: 1655-1661.

Roff, D.A. 2002. Life History Evolution. Sinauer Associates, Sunderland, MA.

Sakai A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, With A. Kimberly, S. Baughman, R.J. Cabin, J.E. Cohen, N.C. Ellstrand, D.E. McCauley, P. O'Neil, I.M. Parker, J.N. Thompson, and S.G. Weller. 2001. The population biology of invasive species. Annual Review of Ecological Systems **32**: 305-332.

Vitousek, P.M., C.M. D'Antonio, L.L. Loope, and R. Westbrooks. 1996. Biological invasions as global environmental change. American Scientist **84**: 218–28.